

## Supplementary Information

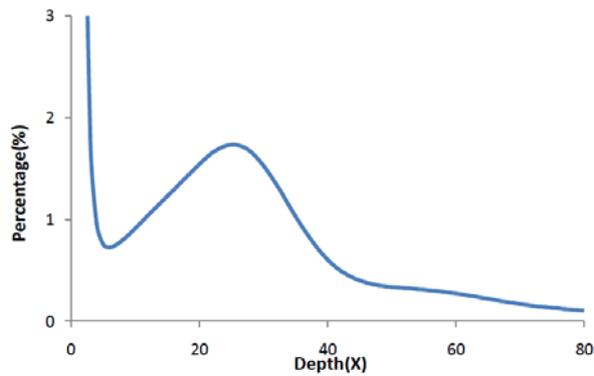
### Minke whale genome and aquatic adaptation in cetaceans

Hyung-Soon Yim<sup>1,24</sup>, Yun Sung Cho<sup>2,24</sup>, Xuanmin Guang<sup>3,24</sup>, Sung Gyun Kang<sup>1,4</sup>, Jae-Yeon Jeong<sup>1,4</sup>, Sun-Shin Cha<sup>1,4,5</sup>, Hyun-Myung Oh<sup>1</sup>, Jae-Hak Lee<sup>1</sup>, Eun Chan Yang<sup>1</sup>, Kae Kyoung Kwon<sup>1,4</sup>, Yun Jae Kim<sup>1</sup>, Tae Wan Kim<sup>1</sup>, Wonduck Kim<sup>1</sup>, Jeong Ho Jeon<sup>1</sup>, Sang-Jin Kim<sup>1,4</sup>, Dong Han Choi<sup>1</sup>, Sungwoong Jho<sup>2</sup>, Hak-Min Kim<sup>2</sup>, Junsu Ko<sup>6</sup>, Hyunmin Kim<sup>6</sup>, Young-Ah Shin<sup>2</sup>, Hyun-Ju Jung<sup>6</sup>, Yuan Zheng<sup>3</sup>, Zhuo Wang<sup>3</sup>, Yan Chen<sup>3</sup>, Ming Chen<sup>3</sup>, Awei Jiang<sup>3</sup>, Erli Li<sup>3</sup>, Shu Zhang<sup>3</sup>, Haolong Hou<sup>7</sup>, Tae Hyung Kim<sup>6</sup>, Lili Yu<sup>3</sup>, Sha Liu<sup>3</sup>, Kung Ahn<sup>6</sup>, Jesse Cooper<sup>6</sup>, Sin-Gi Park<sup>6</sup>, Chang Pyo Hong<sup>6</sup>, Wook Jin<sup>8</sup>, Heui-Soo Kim<sup>9</sup>, Chankyu Park<sup>10</sup>, Kyooyeol Lee<sup>10</sup>, Sung Chun<sup>11</sup>, Phillip Morin<sup>12</sup>, Stephen J. O'Brien<sup>13</sup>, Hang Lee<sup>14</sup>, Jumpei Kimura<sup>15</sup>, Dae Yeon Moon<sup>16</sup>, Andrea Manica<sup>17</sup>, Jeremy Edwards<sup>18</sup>, Byung Chul Kim<sup>2</sup>, Sangsoo Kim<sup>19</sup>, Jun Wang<sup>3,20,21</sup>, Jong Bhak<sup>2,6,22,23</sup>, Hyun Sook Lee<sup>1,4</sup> & Jung-Hyun Lee<sup>1,4</sup>

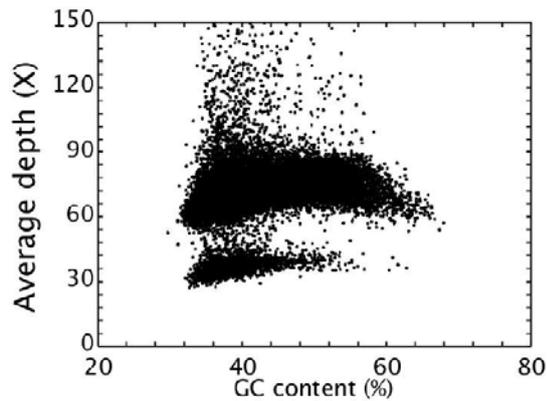
<sup>1</sup>Korea Institute of Ocean Science and Technology, Ansan 426-744, Republic of Korea. <sup>2</sup>Personal Genomics Institute, Genome Research Foundation, Suwon 443-270, Republic of Korea. <sup>3</sup>BGI-Shenzhen, Shenzhen 518083, China. <sup>4</sup>Department of Marine Biotechnology, University of Science and Technology, Daejeon 306-350, Republic of Korea. <sup>5</sup>Ocean Science and Technology School, Korea Maritime University, Pusan 606-791, Republic of Korea. <sup>6</sup>Theragen BiO Institute, TheragenEtex, Suwon 443-270, Republic of Korea. <sup>7</sup>Shaanxi Yulin Energy Group Co. Ltd.; Building 2, Galaxy park, High-tech zone, Yulin, Shaanxi, 719000, China. <sup>8</sup>Department of Molecular Medicine, School of Medicine, Gachon University, Incheon 406-840, Republic of Korea. <sup>9</sup>Department of Biological Sciences, College of Natural Sciences, Pusan National University, Pusan 609-735, Republic of Korea. <sup>10</sup>Laboratory of Genome Biology, Department of Animal Biotechnology, Konkuk University, Seoul 143-701, Republic of Korea. <sup>11</sup>Division of Genetics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA. <sup>12</sup>Protected Resources Division, Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, La Jolla, California 92037, USA. <sup>13</sup>Theodosius Dobzhansky Center for Genome Bioinformatics, St. Petersburg State University, St. Petersburg, Russia. <sup>14</sup>College of Veterinary Medicine, Seoul National University, Seoul 151-742, Republic of Korea. <sup>15</sup>Department of Anatomy and Cell Biology, College of Veterinary Medicine, Seoul National University, Seoul 151-742, Republic of Korea. <sup>16</sup>Marine Biodiversity Institute of Korea (MABIK), Ministry of Ocean and Fisheries, Sejong, Republic of Korea. <sup>17</sup>Evolutionary Ecology Group, Dept. of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK. <sup>18</sup>Molecular Genetics and Microbiology, University of New Mexico Health Sciences Center, USA. <sup>19</sup>School of Systems Biomedical Science, Soongsil University, Seoul 156-743, Republic of Korea. <sup>20</sup>Department of Biology, University of Copenhagen, Copenhagen DK-2200, Denmark. <sup>21</sup>King Abdulaziz University, Jeddah, Saudi Arabia. <sup>22</sup>Program in Nano Science and Technology, Department of Transdisciplinary Studies, Seoul National University, Suwon 443-270, Republic of Korea. <sup>23</sup>Advanced Institutes of Convergence Technology Nano Science and Technology, Suwon 443-270, Republic of Korea. <sup>24</sup>These authors contributed equally to this work.

Correspondence should be addressed to J-H. Lee (jlee@kiost.ac), H.S. Lee (leeh522@kiost.ac), or J. Bhak (jongbhak@genomics.org).

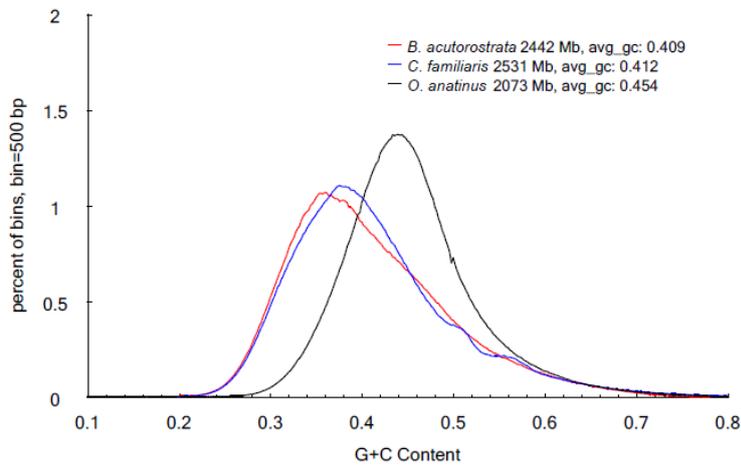
## Supplementary Figures



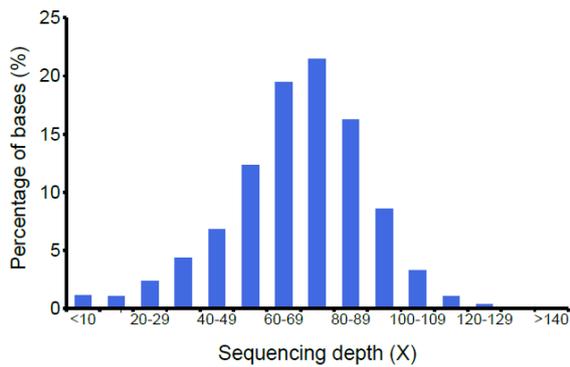
**Supplementary Figure 1 | Estimation of the minke whale genome size based on 17-mer analysis.** The  $x$ -axis represents the depth ( $X$ ) and the  $y$ -axis represents the proportion, i.e., the frequency at that depth divided by the total frequency at all depths.



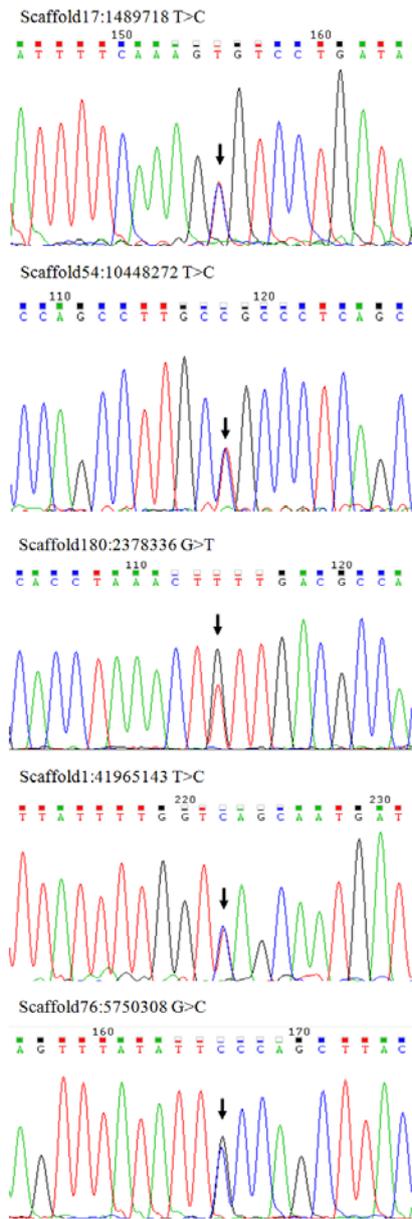
**Supplementary Figure 2 | GC content and sequencing depths.** The  $x$ -axis represents the GC content and the  $y$ -axis represents the average depth.



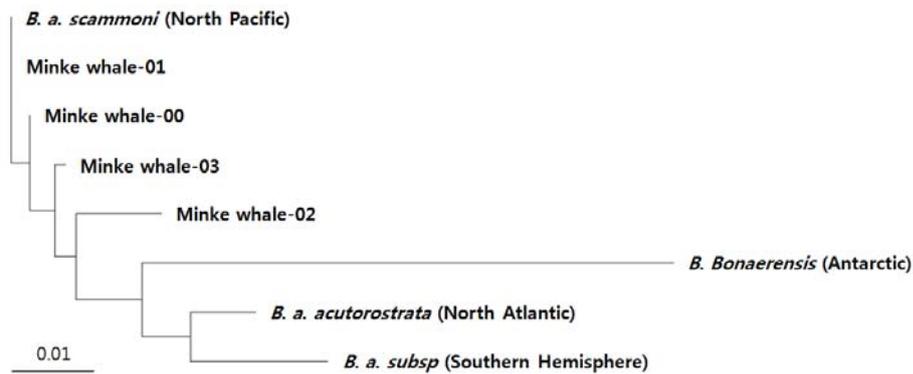
**Supplementary Figure 3 | GC content distribution for the genomes of related species.** The *x*-axis shows the GC content and the *y*-axis represents the ratio of the bin number divided by the total windows.



**Supplementary Figure 4 | Sequencing depth distribution.** The *x*-axis shows the depth and the *y*-axis represents the ratio of the base number divided by the total bases.

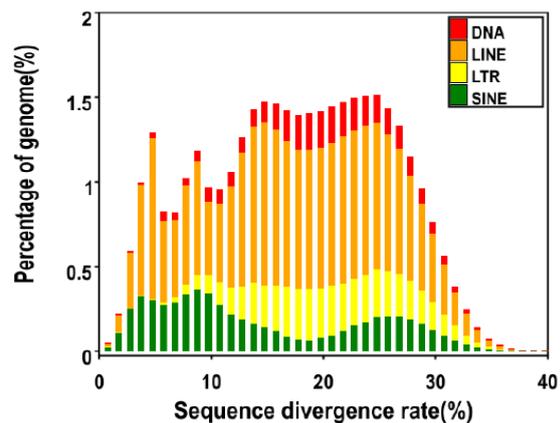


**Supplementary Figure 5 | Examples of the Sanger sequencing results for heterozygous SNVs.** Black arrows are randomly selected heterozygous SNV positions in the minke whale genome.

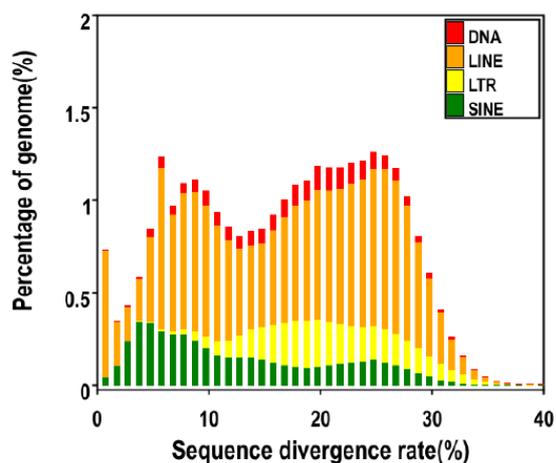


**Supplementary Figure 6 | Species and sub-species identification for four minke whales.** Whole length mtDNA sequences for four minke whales were generated by mapping their reads to the previously reported mitochondrial genome of a North Atlantic minke whale (*B. a. acutorostrata*, accession: NC\_005271.1). Consensus sequences from each of four mtDNA were generated using SAMtools. Control region sequences used in this analysis are EF113709.1 (*B. bonaerensis*), AY878077.1 (*B. a. scammoni*), X87773.1 (*B. a. acutorostrata*), and EU285375.1 (*B. a. subsp.*). All the four minke whales were identified as *B. a. scammoni* (1,000 times bootstrapping, data not shown).

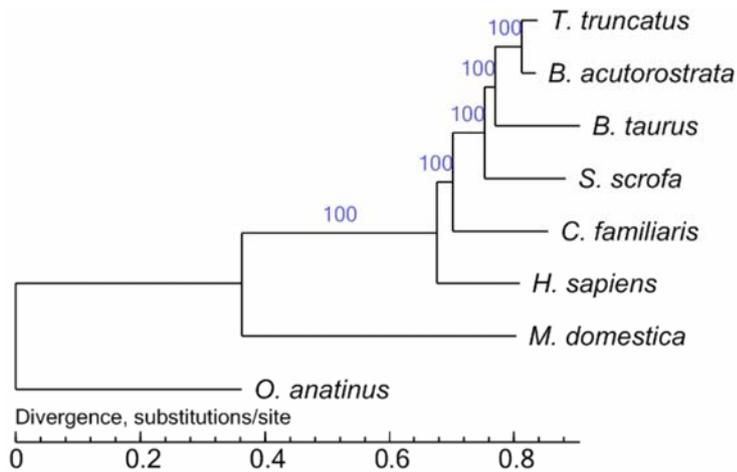




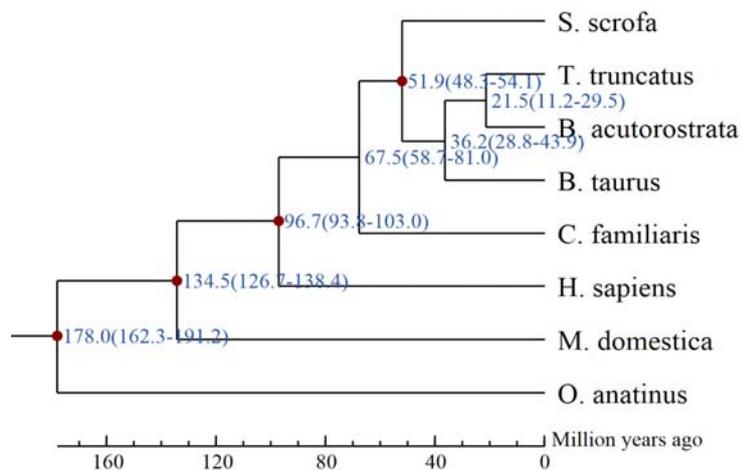
**Supplementary Figure 8 | Distributions of the divergence rates for each TE type.** The divergence rate was calculated for each TE elements identified in the genome using the consensus sequences in Repbase.



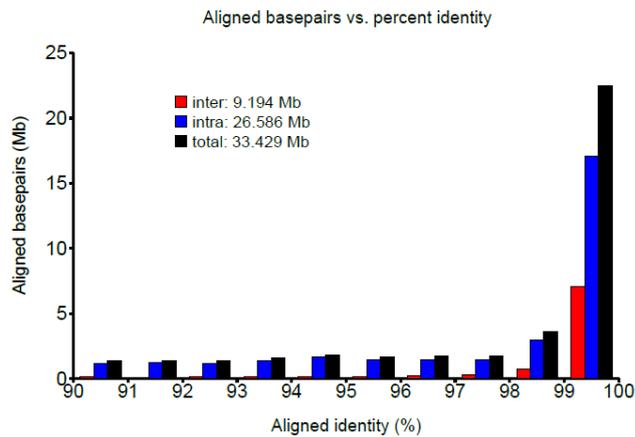
**Supplementary Figure 9 | Distributions of the divergence rates for each TE type by *de novo* method.** The divergence rate was calculated for each TE elements identified in the genome using the consensus sequences in the predicted TE library by the *de novo* method.



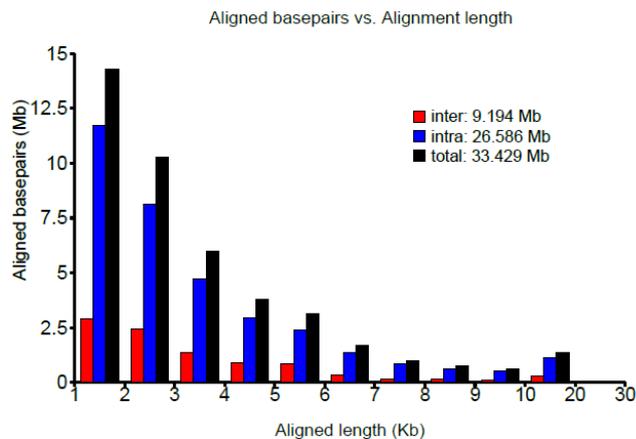
**Supplementary Figure 10 | Phylogenetic tree constructed for orthologous genes with four-fold degeneration sites using the maximum likelihood method.** Branch lengths represent the neutral divergence rates. Blue numbers on the branches represent aLRT values (the default values for assessing branch reliability in PhyML), which illustrate the reliability of branches calculated by PhyML.



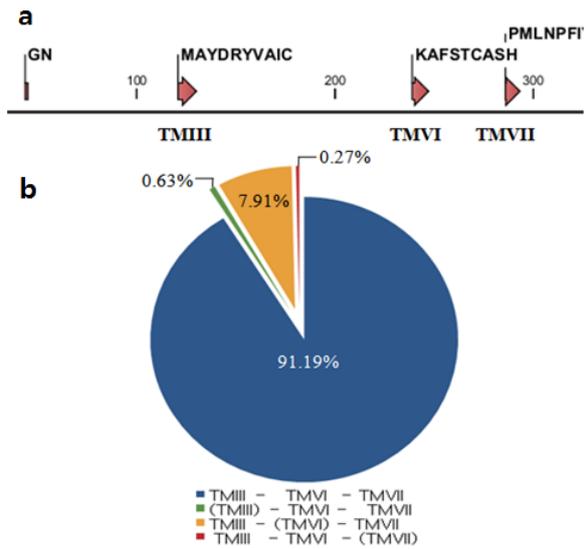
**Supplementary Figure 11 | Estimation of divergence time and substitution rate.** The blue numbers on the nodes represent the divergence time from the present (MYA).



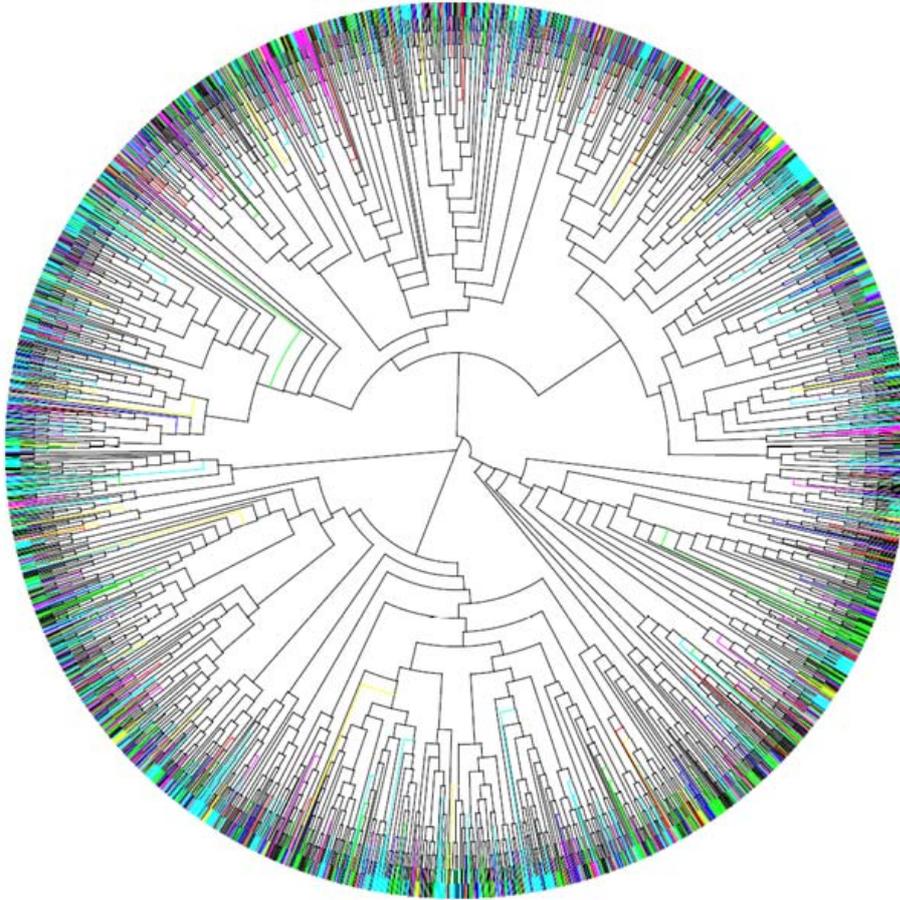
**Supplementary Figure 12 | Distribution of the percent identity of the segmental duplications in the minke whale genome.** The horizontal axis represents the percent identity, and the vertical axis represents the total length of these segmental duplications in each window of percent identity. Inter, intra, and total indicate segmental duplications within scaffolds, between scaffolds, and both within and between scaffolds, respectively.



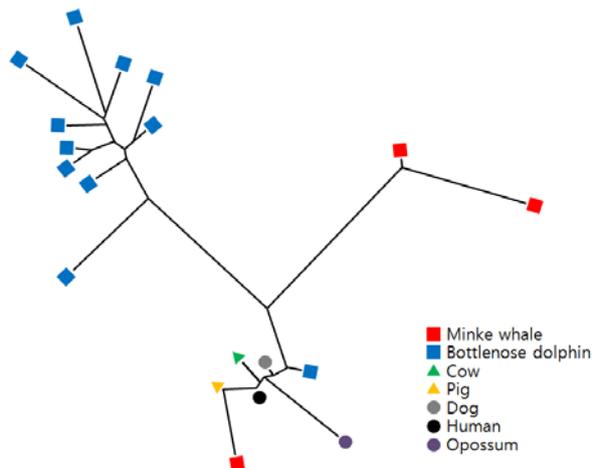
**Supplementary Figure 13 | Distribution of the alignment length of the segmental duplications in the minke whale genome.** The horizontal axis represents the length of the segmental duplications, and the vertical axis represents the total length of these segmental duplications in each window of alignment length. Inter, intra, and totals indicate segmental duplications within scaffolds, between scaffolds, and both within and between scaffolds, respectively.



**Supplementary Figure 14 | Conserved olfactory receptor (OR)-specific motifs used to identify OR genes in the minke whale genome and the frequency of sequences with or without these motifs. a.** The amino acid sequences of the OR-specific motifs are shown. The numbers indicate the positions of amino acids. TM, transmembrane domain. **b.** Proportional distribution of the 134 functional and pseudo-OR amino acid sequences identified based on their OR motif-containing patterns. The motifs within parentheses were absent. The absence or presence of the GN motif is not indicated.



**Supplementary Figure 15 | Contraction of the olfactory receptor gene family in the whale lineage.** The olfactory receptor protein sequences were derived from contracted orthologous gene families in the whale lineage, except for the gene families with no olfactory receptor genes in the whale lineage. The red, green, black, blue, magenta, cyan, and yellow nodes are olfactory receptor genes in the whale lineage, cow, pig, dog, human, opossum, and platypus, respectively. Compared with other mammalian species, the minke whale and bottlenose dolphin had evolutionarily contracted olfactory receptor genes (red), whereas the olfactory receptor genes of the opossum (cyan) were the most highly expanded.



**Supplementary Figure 16 | Expansion of the *O*-GlcNac transferase (*OGT*) gene family in the whale lineage.** *OGT* was expanded in the minke whale (3 copies) and bottlenose dolphin (11 copies).

minke_whale	NCQNEEILNSLKYVRPGGGFQPTFTLIQKCDVNGQNEHPVFIYLKDKLPYPYDDPFSLMT	81
fin_whale	NCQNEEILNSLKYVRPGGGFQPTFTLIQKCDVNGQNEHPVFIYLKDKLPYPYDDPFSLMT	81
bottlenose_dolphin	NCQNEEILNSLKYVRPGGGFQPTFTLIQKCDVNGQNEHPVFIYLKDKLPYPYDDPFSLMT	130
finless_porpoise	NCQNEEILNSLKYVRPGGGFQPTFTLIQKCDVNGQNEHPVFIYLKDKLPYPYDDPFSLMT	130
cow	NCQNEEILNSLKYVRPGGGFQPTFTLVQKCDVNGQNEHPVFIYLKDKLPYPYDDPFSLMT	91
pig	NCQNEEILNSLKYVRPGGGFQPTFTLIQKCDVNGQNEHPVFIYLKDKLPYPYDDPFSLMT	95
dog	NCQNEEILNSLKYVRPGGGFQPTFTLVQKCEVNGQNEHPVFIYLKDKLPYPYDDPFLSMT	94
human	NCQNEEILNSLKYVRPGGGFQPTFTLVQKCEVNGQNEHPVFIYLKDKLPYPYDDPFSLMT	135
mouse	NCQNEEILNSLKYVRPGGGYQPTFSLTQKCDVNGQNEHPVFIYLKDKLPYPYDDPFSLMT	135
	*****:****:* **:******:*****:****	
minke_whale	DPKFIIWSPVRSVSDVSNFEKFLIGPEGEFFRRYSRTFPTINIEPDIKRLKVAI	136
fin_whale	DPKFIIWSPVRSVSDVSNFEKFLIGPEGEFFRRYSRTFPTINIEPDIKRLKVAI	136
bottlenose_dolphin	DPKFIIWSPVRSVSDVSNFEKFLIGPEGEFFRRYSRTFPTINIEPDIKRLKVAI	185
finless_porpoise	DPKFIIWSPVRSVSDVSNFEKFLIGPEGEFFRRYSRTFPTINIEPDIKRLKVAI	185
cow	DPKFIIWSPVRSVSDVSNFEKFLIGPEGEFFRRYSRTFQTINIEPDIKRLKVAI	146
pig	DPKFIIWSPVRSVSDVSNFEKFLIGPEGEFFRRYSRTFPTINIEPDIKRLKVAI	150
dog	DPKFIIWSPVRSVSDVSNFEKFLIGPEGEFFRRYSRTFPTINIEPDIKRLKVAI	148
human	DPKLIIWSPVRSVSDVSNFEKFLIGPEGEFFRRYSRTFPTINIEPDIKRLKVAI	190
mouse	DPKLIIWSPVRSVSDVSNFEKFLIGPEGEFFRRYSRSFQTINIEPDIKRLKVAI	190
	**:******:****:* **:******:*****:****	

**Supplementary Figure 17 | Cetacean-specific amino acid changes in the *GPX2* gene.** Red rectangles indicate cetacean-specific (4 minke whale, 1 fin whale, 2 bottlenose dolphin, and 1 finless porpoise) amino acid changes.



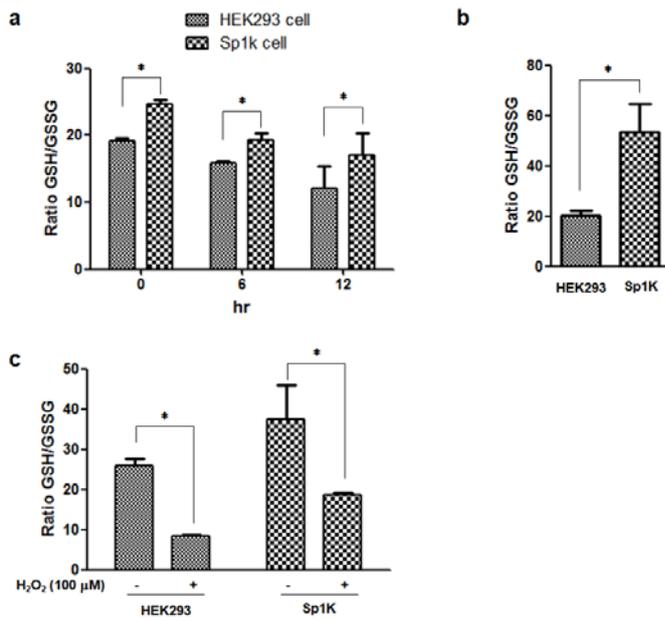


```

minke_whale      MAKGFYISKALGILGILLGMAAATI IALSIVYAQEKKNNAVDPT-----SPATT---TITAATTLQSKPWNRRLPTLLPDSYKVTLRPYLTPDENGLYIFKGTSVIRFICK 108
fin_whale       MAKGFYISKALGILGILLGMAAATI IALSIVYAQEKKNNAVDPT-----SPATT---TITAATTLQSKPWNRRLPTLLPDSYKVTLRPYLTPDENGLYIFKGTSVIRFICK 108
bottlenose_dolphin  MAKGFYISKALGILGILLGMAAATI IALSIVYAQEKKNNAVDPT-----SPATT---TITAATTLQSKPWNRRLPTLLPDSYKVTLRPYLTPDENGLYIFKGTSVIRFICK 110
finless_porpoise  MAKGFYISKALGILGILLGMAAATI IALSIVYAQEKKNNAVDPT-----SPATT---TITAATTLQSKPWNRRLPTLLPDSYKVTLRPYLTPDENGLYIFKGTSVIRFICK 110
cow             MAKGFYISKALGILGILLGMAAATI IALSIVYAQEKKNNAVDPT-----SPATT---TITAATTLQSKPWNRRLPTLLPDSYKVTLRPYLTPDENGLYIFKGTSVIRFICK 111
pig            MAKGFYISKALGILGILLGMAAATI IALSIVYAQEKKNNAVDPT-----SPATT---TITAATTLQSKPWNRRLPTLLPDSYKVTLRPYLTPDENGLYIFKGTSVIRFICK 110
dog            MAKGFYISKALGILGILLGMAAATI IALSIVYAQEKKNNAVDPT-----SPATT---TITAATTLQSKPWNRRLPTLLPDSYKVTLRPYLTPDENGLYIFKGTSVIRFICK 110
human          MAKGFYISKALGILGILLGMAAATI IALSIVYAQEKKNNAVDPT-----SPATT---TITAATTLQSKPWNRRLPTLLPDSYKVTLRPYLTPDENGLYIFKGTSVIRFICK 114
mouse          MAKGFYISKALGILGILLGMAAATI IALSIVYAQEKKNNAVDPT-----SPATT---TITAATTLQSKPWNRRLPTLLPDSYKVTLRPYLTPDENGLYIFKGTSVIRFICK 114
*****:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*
minke_whale      ITLIYPKELTALSNMPPKGPSVPLAEDPKWVTEFQTPVMSTYLLAYIVSEFTSVETSAPNSVQVRIWARPNATAQNHGIYALNVTGPILNFFAHDTPYPLPKSDQIALPDFNAGAM 348
fin_whale       ITLIYPKELTALSNMPPKGPSVPLAEDPKWVTEFQTPVMSTYLLAYIVSEFTSVETSAPNSVQVRIWARPNATAQNHGIYALNVTGPILNFFAHDTPYPLPKSDQIALPDFNAGAM 348
bottlenose_dolphin  ITLIYPNGFTALSNMPPKGPSVPLAEDPKWVTEFQTPVMSTYLLAYIVCEFTSVKTMAPNNVQVGIWARPNATAQNHGIYALNVTGPILNFFAHDTPYPLPKSDQIALPDFNAGAM 350
finless_porpoise  ITLIHPNGFTALSNMPPKGPSVPLAEDPKWVTEFQTPVMSTYLLAYIVCEFTSVKTMAPNNVQVGIWARPNATAQNHGIYALNVTGPILNFFAHDTPYPLPKSDQIALPDFNAGAM 350
cow             ITLIHPKDLTALSNMPPKGPSVPLAEDPKWVTEFQTPVMSTYLLAYIVSEFTSVESVAPNDVQIRIWARPNATAQNHGIYALNVTGPILNFFAHDTPYPLPKSDQIALPDFNAGAM 351
pig            ITLIHPNLTALSNMPPKGPSVPLAEDPKWVTEFQTPVMSTYLLAYIVSEFTSVESVAPNDVQIRIWARPNATAQNHGIYALNVTGPILNFFAHDTPYPLPKSDQIALPDFNAGAM 349
dog            ITLIHPNLTALSNMPPKGPSVPLAEDPKWVTEFQTPVMSTYLLAYIVSEFTSVESVAPNDVQIRIWARPNATAQNHGIYALNVTGPILNFFAHDTPYPLPKSDQIALPDFNAGAM 360
human          ITLIHPKDLTALSNMPPKGPSVPLAEDPKWVTEFQTPVMSTYLLAYIVSEFTSVESVAPNDVQIRIWARPNATAQNHGIYALNVTGPILNFFAHDTPYPLPKSDQIALPDFNAGAM 354
mouse          ITLIYPNLTALSNMPPKGPSVPLAEDPKWVTEFQTPVMSTYLLAYIVSEFTSVESVAPNDVQIRIWARPNATAQNHGIYALNVTGPILNFFAHDTPYPLPKSDQIALPDFNAGAM 353
*****:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*
minke_whale      ENWGLVITYRENSLLFDPOSSSSNKERVVTVIAHELAHQWFGNLVTLAWNDLWLNDFGASVVEYLADYAEPTWNLKDLMPNDYRMAVDALASSHPLTPAEVNTPAQISEMFD 468
fin_whale       ENWGLVITYRENSLLFDPOSSSSNKERVVTVIAHELAHQWFGNLVTLAWNDLWLNDFGASVVEYLADYAEPTWNLKDLMPNDYRMAVDALASSHPLTPAEVNTPAQISEMFD 468
bottlenose_dolphin  ENWGLVITYRENSLLFDPOSSSSNKERVVTVIAHELAHQWFGNLVTLAWNDLWLNDFGASVVEYLADYAEPTWNLKDLMPNDYRMAVDALASSHPLTPAEVNTPAQISEMFD 470
finless_porpoise  ENWGLVITYRENSLLFDPOSSSSNKERVVTVIAHELAHQWFGNLVTLAWNDLWLNDFGASVVEYLADYAEPTWNLKDLMPNDYRMAVDALASSHPLTPAEVNTPAQISEMFD 470
cow             ENWGLVITYRENSLLFDPOSSSSNKERVVTVIAHELAHQWFGNLVTLAWNDLWLNDFGASVVEYLADYAEPTWNLKDLMPNDYRMAVDALASSHPLTPAEVNTPAQISEMFD 471
pig            ENWGLVITYRENSLLFDPOSSSSNKERVVTVIAHELAHQWFGNLVTLAWNDLWLNDFGASVVEYLADYAEPTWNLKDLMPNDYRMAVDALASSHPLTPAEVNTPAQISEMFD 471
dog            ENWGLVITYRENSLLFDPOSSSSNKERVVTVIAHELAHQWFGNLVTLAWNDLWLNDFGASVVEYLADYAEPTWNLKDLMPNDYRMAVDALASSHPLTPAEVNTPAQISEMFD 469
human          ENWGLVITYRENSLLFDPOSSSSNKERVVTVIAHELAHQWFGNLVTLAWNDLWLNDFGASVVEYLADYAEPTWNLKDLMPNDYRMAVDALASSHPLTPAEVNTPAQISEMFD 474
mouse          ENWGLVITYRENSLLFDPOSSSSNKERVVTVIAHELAHQWFGNLVTLAWNDLWLNDFGASVVEYLADYAEPTWNLKDLMPNDYRMAVDALASSHPLTPAEVNTPAQISEMFD 473
*****:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*
minke_whale      ISYKVLPSPRAGPTEVGMVREGASVIRMLSEFLTEDFLKGLASYLHAFAYQNTTYLDLWEHLQAVNNQSAIKLPGTVRDMRWTLQMGFPVITVDTKGTSIQKHFLDPSTV 588
fin_whale       ISYKVLPSPRAGPTEVGMVREGASVIRMLSEFLTEDFLKGLASYLHAFAYQNTTYLDLWEHLQAVNNQSAIKLPGTVRDMRWTLQMGFPVITVDTKGTSIQKHFLDPSTV 588
bottlenose_dolphin  ISYS-----KGASVIRMLSEFLTEDFLKGLASYLHAFAYQNTTYLDLWEHLQAVNNQSAIKLPGTVRDMRWTLQMGFPVITVDTKGTSIQKHFLDPSTV 572
finless_porpoise  ISYS-----KGASVIRMLSEFLTEDFLKGLASYLHAFAYQNTTYLDLWEHLQAVNNQSAIKLPGTVRDMRWTLQMGFPVITVDTKGTSIQKHFLDPSTV 572
cow             ISYS-----KGASVIRMLSNFLTEDFLKGLASYLQTFAYQNTTYLDLWEHLQAVNNQSAIKLPGTVRDMRWTLQMGFPVITVDTKGTSIQKHFLDPSTV 573
pig            ISYS-----KGASVIRMLSNFLTEDFLKGLASYLHAFAYQNTTYLDLWEHLQAVNNQSAIKLPGTVRDMRWTLQMGFPVITVDTKGTSIQKHFLDPSTV 571
dog            ISYS-----KGASVIRMLSNFLTEDFLKGLASYLHAFAYQNTTYLDLWEHLQAVNNQSAIKLPGTVRDMRWTLQMGFPVITVDTKGTSIQKHFLDPSTV 582
human          ISYS-----KGASVIRMLSNFLTEDFLKGLASYLHAFAYQNTTYLDLWEHLQAVNNQSAIKLPGTVRDMRWTLQMGFPVITVDTKGTSIQKHFLDPSTV 575
mouse          ITYS-----KGASVIRMLSNFLTEDFLKGLSSYLHTYQNSNTYLDLWEHLQAVNNQSAIKLPGTVRDMRWTLQMGFPVITVDTKGTSIQKHFLDPSTV 575
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*
minke_whale      WQAASSLRYFKLMFDRSEVYGMKKNYLRQVPEPLFLYFENLTKNWTVPESLMDQYNEINAI STACSNGLPKCEELAKNLYNQWENPQNNPIHPNLRSTIYCNAIAGQKQEWDFAWR 828
fin_whale       WQAASSLRYFKLMFDRSEVYGMKKNYLRQVPEPLFLYFENLTKNWTVPESLMDQYNEINAI STACSNGLPKCEELAKNLYNQWENPQNNPIHPNLRSTIYCNAIAGQKQEWDFAWR 828
bottlenose_dolphin  WQAASSLRYFKLMFDRSEVYGMKKNYLRQVPEPLFLYFENLTKNWTVPESLMDQYNEINAI STACSNGLPKCEELAKNLYNQWENPQNNPIHPNLRSTIYCNAIAGQKQEWDFAWR 811
finless_porpoise  WQAASSLRYFKLMFDRSEVYGMKKNYLRQVPEPLFLYFENLTKNWTVPESLMDQYNEINAI STACSNGLPKCEELAKNLYNQWENPQNNPIHPNLRSTIYCNAIAGQKQEWDFAWR 811
cow             WQAAVSSLYFKLMFDRSEVYGMKKNYLRQVPEPLFLYFENLTKNWTVPESLMDQYNEINAI STACSNGLPKCEELAKNLYNQWENPQNNPIHPNLRSTIYCNAIAGQKQEWDFAWN 813
pig            WQAAVSSLYFKLMFDRSEVYGMKKNYLRQVPEPLFLYFENLTKNWTVPESLMDQYNEINAI STACSNGLPKCEELAKNLYNQWENPQNNPIHPNLRSTIYCNAIAGQKQEWDFAWN 811
dog            WEAAVSSLYFKLMFDRSEVYGMKKNYLRQVPEPLFLYFENLTKNWTVPESLMDQYNEINAI STACSNGLPKCEELAKNLYNQWENPQNNPIHPNLRSTIYCNAIAGQKQEWDFAWN 822
human          WEAAVSSLYFKLMFDRSEVYGMKKNYLRQVPEPLFLYFENLTKNWTVPESLMDQYNEINAI STACSNGLPKCEELAKNLYNQWENPQNNPIHPNLRSTIYCNAIAGQKQEWDFAWN 814
mouse          WQAAVSSLYFKLMFDRSEVYGMKKNYLRQVPEPLFLYFENLTKNWTVPESLMDQYNEINAI STACSNGLPKCEELAKNLYNQWENPQNNPIHPNLRSTIYCNAIAGQKQEWDFAWN 813
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

```

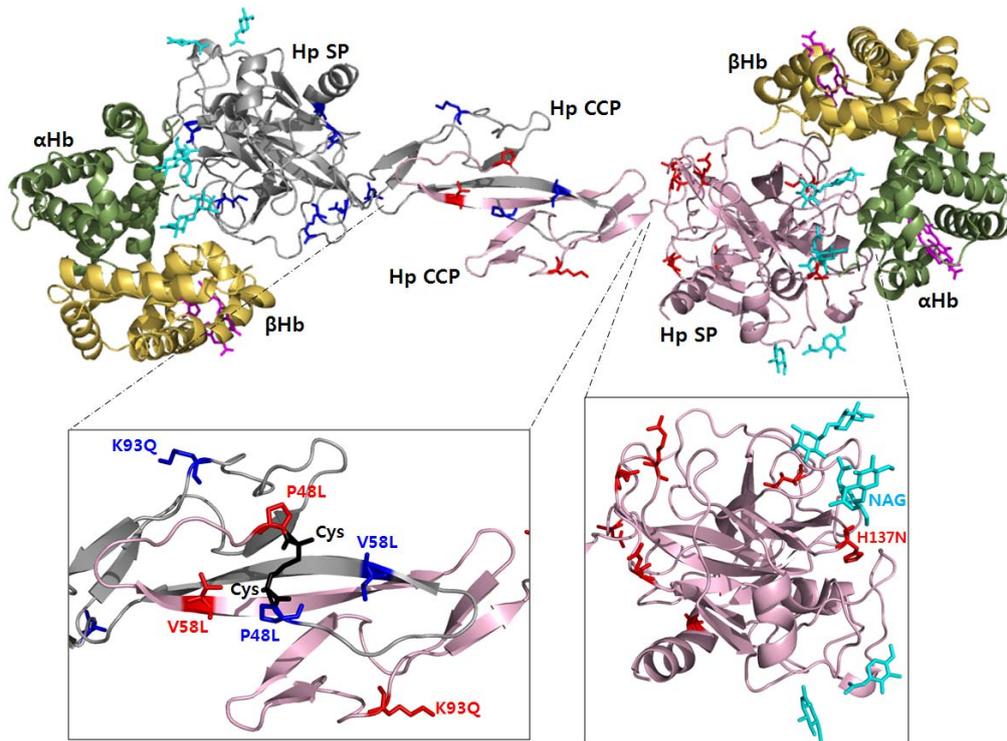
**Supplementary Figure 23 | Cetacean-specific amino acid changes in the ANPEP gene.**  
 Red rectangles indicate cetacean-specific (4 minke whale, 1 fin whale, 2 bottlenose dolphin, and 1 finless porpoise) amino acid changes.



**Supplementary Figure 24 | Relationship between the GSH/GSSG redox status and hypoxia/H<sub>2</sub>O<sub>2</sub> treatment in HEK293 and Sp1k cells.** **a.** HEK293 and Sp1k cells were exposed to hypoxia for 0, 6, and 12 h. The levels of GSH/GSSG ratio were determined. \**P* < 0.05 according to the Student's *t*-test. **b.** HEK293 and Sp1k cells were exposed to hypoxia for 12 h and returned to atmospheric oxygen for 5 min. The GSH/GSSG ratios were determined. \**P* < 0.05 according to the Student's *t*-test. **c.** HEK293 and Sp1k cells were exposed to normoxia with or without H<sub>2</sub>O<sub>2</sub> (100 μM) for 1 h. The GSH/GSSG ratios were determined. \**P* < 0.05 according to the Student's *t*-test.

minke_whale	-----DDSCIKPPEIPNGYLEHLVRYRC	66
fin_whale	-----DDSCIKPPEIPNGYLEHLVRYRC	66
bottlenose_dolphin	-----DDSCIKPPEIANGYLEHLVRYRC	65
finless_porpoise	-----DDSCIKPPEIANGYLEHLVRYRC	65
cow	YQCDKYYKLHAGNGVYTFNNKQWINKDIGQQLPECEEDDSCPEPKIENGYEYLVRYQC	106
pig	-----DDSCIKPPEIPKGYEYHVRVYHC	52
human	-----DDGCKPPEIAHGYEYHSVRYQC	52
mouse	-----DDSCIKPPEIANGYLEHLVRYRC	52
minke_whale	-----EKQWLNKDLGELPECEAACGPKHPVQVQR I IGGSLDAKGS	126
fin_whale	-----EKQWLNKDLGELPECEAACGPKHPVQVQR I IGGSLDAKGS	126
bottlenose_dolphin	-----EKQWLNKDLGELPECEAVCGPKHPVQVQR I IGGSLDAKGS	125
finless_porpoise	-----EKQWLNKDLGELPECEAVCGPKHPVQVQR I IGGSLDAKGS	125
cow	-----KKQWLNKDLGELPECEAVCGPKHPVQVQR I IGGSLDAKGS	167
pig	-----NKQWLNKDLGELPECEAVCGPKHPVQVQR I IGGSLDAKGS	113
human	NYYKLRTEGDGYYTLNNEKQWINKAVGDKLPECEAVCGPKHPVQVQR I IGGSLDAKGS	172
mouse	-----EKQWLNKDLGELPECEAVCGPKHPVQVQR I IGGSLDAKGS	113
minke_whale	FPWQAKMVSNNLTSGATL INEQWLLTTAKNLF LGHDNTKAKDI APTLRL YVGRKQVLE	186
fin_whale	FPWQAKMVSNNLTSGATL INEQWLLTTAKNLF LGHDNTKAKDI APTLRL YVGRKQVLE	186
bottlenose_dolphin	FPWQAKMVSNNLTSGATL INEQWLLTTAKNLF LGHDNTKAKDI APTLRL YVGRKQVLE	185
finless_porpoise	FPWQAKMVSNNLTSGATL INEQWLLTTAKNLF LGHDNTKAKDI APTLRL YVGRKQVLE	185
cow	FPWQAKMVSNNLTSGATL INERWLLTTAKNLF LGHSDKAKDI TPTLRL YVGRKQVLE	227
pig	FPWQAKMVSNNLTSGATL INEQWLLTTAKNLF LGHDNTKAKDI APTLRL YVGRKQVLE	173
human	FPWQAKMVSNNLTSGATL INEQWLLTTAKNLF LNHSNATAKDI APTLRL YVGRKQVLE	232
mouse	FPWQAKMVSNNLTSGATL ISDQWLLTTAKNLF LNHSNATAKDI TPTLRL YVGRKQVLE	173
minke_whale	IEKVLHFDYSEVDI GLIKLREKVPTEIEMPI CLPSKDYVEYGRVGYVSGWGRNANLIF	246
fin_whale	IEKVLHFDYSEVDI GLIKLREKVPTEIEMPI CLPSKDYVEYGRVGYVSGWGRNANLIF	246
bottlenose_dolphin	IEKVLHFDYSEVDI GLIKLREKVPTEIEMPI CLPSKDYVEYGRVGYVSGWGRNANLIF	245
finless_porpoise	IEKVLHFDYSEVDI GLIKLREKVPTEIEMPI CLPSKDYVEYGRVGYVSGWGRNANLIF	245
cow	VEKVLHFDHSEVDI GLIKLREKVPTEIEMPI CLPSKDYVEYGRVGYVSGWGRNANLIF	287
pig	IEKVLHFDYSEVDI GLIKLREKVPTEIEMPI CLPSKDYVEYGRVGYVSGWGRNANLIF	233
human	IEKVLHFDYSEVDI GLIKLREKVPTEIEMPI CLPSKDYVEYGRVGYVSGWGRNANLIF	292
mouse	IEKVLHFDYSEVDI GLIKLREKVPTEIEMPI CLPSKDYVEYGRVGYVSGWGRNANLIF	233
minke_whale	TEHLKYVMLPVADQDCVHYEGSTVPEKKTSPVGVQPI LNEHTFCAGLSKYQEDTCY	306
fin_whale	TEHLKYVMLPVADQDCVHYEGSTVPEKKTSPVGVQPI LNEHTFCAGLSKYQEDTCY	306
bottlenose_dolphin	TEHLKYVMLPVADQDCVHYEGSTVPEKKTSPVGVQPI LNEHTFCAGLSKYQEDTCY	305
finless_porpoise	TEHLKYVMLPVADQDCVHYEGSTVPEKKTSPVGVQPI LNEHTFCAGLSKYQEDTCY	305
cow	TEHLKYVMLPVADQDCVHYEG IDAPKNKTAKSPVGVQPI LNEHTFCAGLSKYQEDTCY	347
pig	TEHLKYVMLPVADQDCVHYEGSTVPEKKTSPVGVQPI LNEHTFCAGLSKYQEDTCY	293
human	TDHLKYVMLPVADQDCVHYEGSTVPEKKTSPVGVQPI LNEHTFCAGLSKYQEDTCY	352
mouse	TDHLKYVMLPVADQDCVHYENSTVPEKKNLTPVGVQPI LNEHTFCAGLSKYQEDTCY	293
minke_whale	GDAGSAFAIHDEADDTWYAAG ILSFDKSCATAEYGVYVKVSSILDWVQKTIADN	360
fin_whale	GDAGSAFAIHDEADDTWYAAG ILSFDKSCATAEYGVYVKVSSILDWVQKTIADN	360
bottlenose_dolphin	GDAGSAFAIHDEADDTWYAAG ILSFDKSCATAEYGVYVKVSSILDWVQKTIADN	359
finless_porpoise	GDAGSAFAIHDEADDTWYAAG ILSFDKSCATAEYGVYVKVSSILDWVQKTIADN	359
cow	GDAGSAFAIHDEADDTWYAAG ILSFDKSCAVAEYGVYVKVSSILDWVQKTIADN	401
pig	GDAGSAFAIHDEADDTWYAAG ILSFDKSCRTAEYGVYVKVSSILDWVQKTIADN	347
human	GDAGSAFAIHDEADDTWYAAG ILSFDKSCAVAEYGVYVKVSSILDWVQKTIADN	406
mouse	GDAGSAFAIHDEADDTWYAAG ILSFDKSCAVAEYGVYVKVSSILDWVQKTIADN	347

**Supplementary Figure 25 | Unique amino acid changes in haptoglobin.** Ten cetacean-specific amino acid changes (four minke whales, a fin whale, two (one reference and one re-sequenced) bottlenose dolphins, and a finless porpoise) were found in the haptoglobin gene. Significant side group changes are shown in the red box, and similar side-group changes are shown in yellow. The green inverted triangle indicates a cysteine residue forming an internal disulfide bond between two haptoglobins, while the other inverted triangles indicate cysteine residues forming an intra-disulfide bond. The black triangles indicate *N*-linked glycosylation sites.



**Supplementary Figure 26 | Structural positions of whale-specific amino acid changes in haptoglobin.** Haptoglobin molecules (Hp, complement control protein (CCP) domain and serine protease (SP) domain) are shown in gray and pink. The hemoglobin (Hb)  $\alpha$ - and  $\beta$ -subunits are green and yellow, respectively. Glycosylations (*N*-acetyl-*D*-glucosamine (NAG)) are shown as cyan sticks and heme groups are shown as magenta sticks. The 10 whale-specific amino acid changes are shown as red and blue sticks. The black sticks indicate a disulfide bridge between two Hp CCPs. The structure of the haptoglobin-hemoglobin complex was obtained from protein data bank (PDBID: 4F4O, organism: *Sus scrofa*).







```

minke_whale      KDVLQEPKTEDIVAVQAKATLYRSCINESAIDSRGGDPLLRLLPDYDWPVTEENWEQTYG-SWTAEKSILAQLNSKYGKKVIINFFVGTDDKNSTNHVIHDQPRLGLPSRDYYECTGIY 237
fin_whale       KDVLQEPKTEDIVAVQAKATLYRSCINESAIDSRGGDPLLRLLPDYDWPVTEENWQTYG-SWTAEKSILAQLNSKYGKKVIINFFVGTDDKNSTNHVIHDQPRLGLPSRDYYECTGIY 237
bottlenose_dolphin KDVLQEPKTEDIVAVQAKATLYRSCINESAD-SRGGVPLLRLLPDYDWPVTEENWQTYGSSWTAEKSILAQLNSKYGKKVIINFFVGTDDKNSTNHVIHDQPRLGLPSRDYYECTGIY 237
finless_porpoise KDVLQEPKTEDIVAVQAKATLYRSCINESAD-SRGGVPLLRLLPDYDWPVTEENWDQTYSSWTAEKSILAQLNSKYGKKVIINFFVGTDDKNSTNHVIHDQPRLGLPSRDYYECTGIY 237
cow             KDVLQEPKTEDIIVAQKAKMLYRSCINETAIESRGGEPLLRLLPDYDWPVTEENWEQTYGSSWTAEKSILAQLNSKYGKKVIINFFVGTDDKNSTNHVIHDQPRLGLPSRDYYECTGIY 238
pig            KDVLQEPKTEDIIVAVQAKATLYRSCINESAIDSRGGEPLLRLLPDYDWPVTEENWEQTYGSSWTAEKSILAQLNSKYGKKVIINFFVGTDDKNSTNHVIHDQPRLGLPSRDYYECTGIY 238
dog            KDVLQEPKTEDIIVAVQAKATLYRSCINESAIDSRGGEPLLRLLPDYDWPVTEENWEQTYGSSWTAEKSILAQLNSKYGKKVIINFFVGTDDKNSTNHVIHDQPRLGLPSRDYYECTGIY 239
human          KDVLQEPKTEDIIVAVQAKALYRSCINESAIDSRGGEPLLRLLPDYDWPVTEENWEQTYGASWTAEKSILAQLNSKYGKKVIINFFVGTDDKNSTNHVIHDQPRLGLPSRDYYECTGIY 238
mouse          KDVLQEPKTEDIIVAVQAKATLYRSCINESAIDSRGGQPLLRLLPDYDWPVTEENWDQTYGTSWTAEKSILAQLNSKYGKKVIINFFVGTDDKNSTNHVIHDQPRLGLPSRDYYECTGIY 238
*****
minke_whale      KEACTAYVDFMISVAKLIRQEKGLPVDEQLSSEMNRVMELEKEIANVKHIFSDTLNNVNHFLFLSLLYRILRATTKSEDRNDPMLLYNKMTLAQIQNNFSLEISGKPFSWSNFTNEIMS 357
fin_whale       KEACTAYVDFMISVAKLIRQEKGLPVDEQLSLEMNRVMELEKEIANVKHIFSDTLNNVNHFLFLSLLYRILRATTKSEDRNDPMLLYNKMTLAQIQNNFSLINGTPFSWSNFTNEIMS 357
bottlenose_dolphin KEACTAYVDFMISVAKLIRQERGLPVDEQLSLEMNRVMELEKEIAN-----ATTKSEDRNDPMLLYNKMTLAQIQNNFSLEISGKXXXXXXXXXXXXX 331
finless_porpoise KEACTAYVDFMISVAKLIRQERGLPVDEQLSLEMNRVMELEKEIAN-----ATTKSEDRNDPMLLYNKMTLAQIQNNFSLEISGKXXXXXXXXXXXXX 331
cow             KEACTAYVDFMISVAKLIRQERGLPIDEQLSLEMNRVMELEKEIAN-----ATTKSEDRNDPMLLYNKMTLAEVQNFSLEISGKPFSWSNFTNEIMS 332
pig            KEACTAYVDFMISVAKLIRQEKGLPVDEQLSLEMNRVMELEKEIAN-----ATTKSEDRNDPMLLYNKMTLAQIQNNFSLEISGKPFSWSNFTNEIMS 332
dog            EEACTAYVDFMISVARLIRQEKGLVIDEQLSLEMNRVMELEKEIAS-----ATTKPEDRNDPMLLYNKMTLAQIQNNFSLEISGKPFSWSNFTNEIMS 333
human          KEACTAYVDFMISVARLIRQERLPIDEQLALEMNRVMELEKEIAN-----ATKPEDRNDPMLLYNKMTLAQIQNNFSLEISGKPFSWLFTNEIMS 332
mouse          KEACTAYVDFMISVARLIRQEQSLPIDEQLSLEMNRVMELEKEIAN-----ATTKPEDRNDPMLLYNKMTLAKLQNNFSLEISGKPFSWSNFTNEIMS 332
:*****

```

**Supplementary Figure 32 | Cetacean-specific amino acid changes in the MME gene.** Red rectangles indicate cetacean-specific (4 minke whale, 1 fin whale, 2 bottlenose dolphin, and 1 finless porpoise) amino acid changes.

```

minke_whale      MKPPAESSGASPSALGNSGLSPGCYSDPDEPEADTPGVGFCLGPACAGDVLDAAPCSAVNYLRWDLSAQQIGELTTLKIETKRVDYHVGSGQELQDVSENTLKALADVESVYTVQR 120
fin_whale       MKPPAESSPGASPSAPGNGLSPGCYSDPDEPEADTPGVGFCLGPACAGDALVAAPCSAVNYLRWDLSAQQIGELTTLKIETKRVDYHVGSGQELQDVSENTLKALADVESVYTVQR 120
bottlenose_dolphin MKPPA-----ACAGDALVAAPCSAVNYLRWDLSAQQIGELTTELIETKRVDYHVGSGQELQDVSENTLKALADVESVYTVQR 79
finless_porpoise MKPPA-----ACAGDALVAAPCSAVNYLRWDLSAQQIGELTTELIETKRVDYHVGSGQELQDVSENTLKALADVESVYTVQR 79
cow             MKPPA-----ACTGDALVAAPCSAVNHLRDLSAQQIAELTTELIETKRVDYHVGQAQELQDVSENTLKALADVESVYTVQR 79
pig            MKPPA-----ACAGDALVAAPCSAVNYLRWDLSAQQIGELTTELIETKRVDYHVGQELQDVSENTLKALADVESVYTVQR 79
dog            MKPPA-----ACVGDALVAAPCSAVNYLRWDLSAQQIEELTTELIETKRVDYHVGQSQKFEDVSESTLKALADVESVYTVQR 79
human          MKPPA-----ACAGMDAAASPCSVNYLRWDLSAQQIEERTRELIETKRVDYHVGQEFEDVSESTLKALADVESVYTVQR 79
mouse          MKPPA-----ACAGDVDAASPASTVNYLRWDLSAQQIRALTTLQIETKRVDYHVGQAQNFEDVSESTLKALADVESVYTVQR 79
*****
minke_whale      NILDFPQHVSPSKDIRTASTEADKKLSEFDVEMSMRQDVYQRIVWLQEKVQDSLPEASRYLERLIRLGRRNGHLPEETQEKIKSIKKLSLLCIDFNKNLNETDTFLPFTREELGGL 240
fin_whale       NILDFPQHVSPSKDIRTASTEADKKLSEFDVEMSMRQDVYQRIVWLQEKVQDSLPEASRYLERLIRLGRRNGHLPEETQEKIKSIKKLSLLCIDFNKNLNETDTFLPFTREELGGL 240
bottlenose_dolphin NILDFPQHVSPSKDIRTASTEADKKLSEFDVEMSMRQDVYQRIVWLQEKVQDSLPEASRYLERLIRLGRRNGHLPEETQEKIKSIKKLSLLCIDFNKNLNETDTFLPFTREELGGL 199
finless_porpoise NILDFPQHVSPSKDIRTASTEADKKLSEFDVEMSMRQDVYQRIVWLQEKVQDSLPEASRYLERLIRLGRRNGHLPEETQEKIKSIKKLSLLCIDFNKNLNETDTFLPFTREELGGL 199
cow             NILDFPQHVSPSKDIRTASTEADKKLSEFDVEMSMRQDVYQRIVWLQEKVQDSLPEASRYLERLIRLGRRNGHLPEETQEKIKSIKKLSLLCIDFNKNLNETDTFLPFTREELGGL 199
pig            NILDFPQHVSPSKDIRTASTEADKKLSEFDVEMSMRQDVYQRIVWLQEKVQDSLPEASRYLERLIRLGRRNGHLPEETQEKIKSIKKLSLLCIDFNKNLNETDTFLPFTREELGGL 199
dog            NILDFPQHVSPSKDIRTASTEADKKLSEFDVEMSMRQDVYQRIVWLQEKVQDSLPEALRYLERLIRLGRRNGHLPEETQEKIKSIKKLSLLCIDFNKNLNETDTFLPFTREELGGL 199
human          NILDFPQHVSPSKDIRTASTEADKKLSEFDVEMSMRQDVYQRIVWLQEKVQDSLPEAARYLERLIRLGRRNGHLPEETQENIKSIKKLSLLCIDFNKNLNETDTFLPFTREELGGL 199
mouse          NILDFPQHVSPSKDIRTASTEADKKLSEFDVEMSMRQDVYQRIVWLQEKVQDSLPEAARYLERLIRLGRRNGHLPEETQEKIKSIKKLSLLCIDFNKNLNETDTFLPFTREELGGL 199
*****
minke_whale      VDQALHTQTPADPAQEYARLQELGVVPATPGTNMPATFGHLAGGYDAQYGYLWSEVYSADMFHTRFKQEGVLSGKVMGYRSCILRPGGSEDASVMLRFLGRDPKQDAFLLSKGLQV 720
fin_whale       VDQALHTQTPADPAQENARLQELGVVPATPGTNMPATFGHLAGGYDAQYGYLWSEVYSADMFHTRFKQEGVLSGKVMGYRSCILRPGGSEDASVMLRFLGRDPKQDAFLLSKGLQV 720
bottlenose_dolphin VDQALHTQTAADPAEEYARLQELGVVPATPGTNMPATFGHLAGGYDAQYGYLWSEVYSADMFHTRFKQEGVLSGKVMGYRSCILRPGGSEDASVMLRFLGRDPKQDAFLLSKGLQV 670
finless_porpoise VDQALHTQTAADPAEEYARLQELGVVPATPGTNMPATFGHLAGGYDAQYGYLWSEVYSADMFHTRFKQEGVLSGKVMGYRSCILRPGGSEDASVMLRFLGRDPKQDAFLLSKGLQV 670
cow             VDQALHTQTAADPAQEYARLQELGVVPATPGTNMPATFGHLAGGYDAQYGYLWSEVYSADMFHTRFKQEGVLSGKVMGYRSCILRPGGSEDASVMLRFLGRDPKQDAFLLSKGLQV 679
pig            VDQALHTQTAADPAEEYARLQELGVVPATPGTNMPATFGHLAGGYDAQYGYLWSEVYSADMFHTRFKQEGVLSGKVMGYRSCILRPGGSEDASVMLRFLGRDPKQDAFLLSKGLQV 678
dog            VDQALHTQTAADPAEEYARLQELGVVPATPGTNMPATFGHLAGGYDAQYGYLWSEVYSADMFHTRFKQEGVLSGKVMGYRSCILRPGGSEDASVMLRFLGRDPKQDAFLLSKGLQV 679
human          VDQALHTQTDADPAEEYARLQELGVVPATPGTNMPATFGHLAGGYDAQYGYLWSEVYSMDMFHTRFKQEGVLSKVMGYRSCILRPGGSEDASAMRRFLGRDPKQDAFLLSKGLQV 679
mouse          VDQVLTQTADPAEEYARLQELGVVPATPGTNMPATFGHLAGGYDAQYGYLWSEVYSMDMFHTRFKQEGVLSKVMGYRSCILRPGGSEDASAMRRFLGRDPKQDAFLLSKGLQV 679
*****

```

**Supplementary Figure 33 | Cetacean-specific amino acid changes in the THOPI gene.** Red rectangles indicate cetacean-specific (4 minke whale, 1 fin whale, 2 bottlenose dolphin, and 1 finless porpoise) amino acid changes.

minke_whale	---NPFYGFYGYHGFGGPPYYSEEMFEQDFEKPKEKDPKTESPATDPSANSTVPEINSHSNAPNPRGSQGGNDTSPTGNSVQGSNTVSNPTAQNSPA	97
fin_whale	---NPFYGFYGYHGFGGPPYYSEEMFEQDFEKPKEKDPKTESPATDPSANSTVPEINSHSNAPNPRGSQGGNDTSPTGNSVQGSNTVSNPTAQNSPA	97
bottlenose_dolphin	EGGNYYLEHD-----GRPPLYSEEMQD--FEK--EKDP--KESPATDHGNSTVETNTQS-----NPRGSQGGNDTSPTGNSVQGSNTVSNPTAQNSPA	276
finless_porpoise	EGGNYYLEHD-----GRPPLYSEEMQD--FEK--EKDP--KESPATDHGNSTVETNTQS-----NPRGSQGGNDTSPTGNSVQGSNTVSNPTAQNSPA	276
cow	EGGNPFYGFYGYHGFGGPPYYSEEMFE--DFEKPKEEDPKTETPATDPSVNSTVPEINSTQGPAPSPRAGQGGNDTSAPAGNQQDPNTVSNPTVQNNPV	296
pig	EGGNPFYGFYGYHGFGGPPYYSEEMFEQDFEKPKEKDPKTESPATDPSVNSTVPEINSTQGPAPSPRAGQGGNDTSAPAGNQQDPNTVSNPTVQNNPV	288
minke_whale	VNVSQQGGPRSQTPRGPSQTNIHENYPNPNIRSFPAQRQWRPTGTVMGHRNRNGPFY*	153
fin_whale	VNVSQQGGPRSQTPRGPSQTNIHENYPNPNIRSFPAQRQWRPTGTVMGHRNRNGPFY*	153
bottlenose_dolphin	VNVSQQGVPRSQTPQGPSQTNIHENYPNPNIRSFPAQRQWRPTGTATGHRNRNGPFYVQQIQRGPRWNSFVLEKQAVRPGYPIYRRAYASTARGNSPNY	376
finless_porpoise	VNVSQQGVPRSQTPQGPSQTNIHENYPNPNIRSFPAQRQWRPTGTAMGHRNRNGPFYVQQIQRGPRWNSFVLEKQAVRPGYPIYRRAYASTARGNSPNY	376
cow	VNVSQQGVPRSQTPWRPSQTNIHENYPNPNIRSFPAQRQWRPTGTSMGNRRNGPFYVQQIQRAPRWNSFVLEKQAVRPGYPIYRRAYASTVRGNYPNY	396
pig	VNVSQQGVPRSQSPWGRPQTIIHENYPNPNIRGFPARRQWRPFGPAMGHRNRNGPFYVQQIQRGPRWNSFVLEKQAVRPGYPIYRRAYASTARGNSPNY	388
minke_whale	-----MRQENPHYPMNTDPKKTIPYNEEDPIDPTGDESFGQSRWGVAESSFKGDPTRVHYESEQYTSN	219
fin_whale	-----MRQENPHYPMNTDPKKTIPYNEEDPIDPTGDESFGQSRWGVAESSFKGDPTRVHYESEQYTSN	219
bottlenose_dolphin	GRVYYPDYNPYDRENPSYPRSNXX	673
finless_porpoise	GRVYYPDYIPYDRENPSYLRSNXX	673
cow	GRVYYDYNPDRRENPSYLRSNWDERNDPPNMGQENPHYPMNTDPKKTIPYNEEGPADPTGDETFPGQTRWVDESNFKTAPTAR--YEGKQYTSN	694
pig	VHYYVDYNPYDRENPSYLRSNWYERDDSPNMGQENPHYPMNTDPKKTIPYNEEDPIDPTGDESFGQSRWDMELSFEDTRVHYESEQYTSN	688
minke_whale	QPKEYLPSLDNPSKPREDFPYGEFYWPSPDENFNSYNTAPTEPRPVESRGYANNVAQGEESTMFPSWNSWGHRIQAQEQKERRPYFNRFWDQPTNLH	319
fin_whale	QPKEYLPSLDNPSKPREDFPYGEFYWPSPDENFNSYNTAPTEPRPVESRGYANNVAQGEESTMFPSWNSWGHRIQAQEQKERRPYFNRFWDQPTNLH	319
bottlenose_dolphin	XX	773
finless_porpoise	XX	773
cow	QPKEYLPSLDNLPKPREYFYGEFYWPSPDENFNSYNTAPTIPLLVENRGIYPTNAVQGEENTMFPSWNSWDMVQVQEQKERRPYFTRTFWQPTNLP	794
pig	QPKEYLPSLDNPSKPREDFLYGEFYVPTPEENFNSYNTAPTISYSPYESRGIYANNVAQGEESTMFPSWNSWDPRIQAQEQKERRPYFNRFWDQPTNLY	788
minke_whale	RAPASSFYQKENKPYSSNSPAGLQKNPIWHEGENLYGMQITRLNSPEREHLAFPDLPYSPYAVQKGAHLFHLQRGPCAGGSTGPKDNPLALQDYTP	419
fin_whale	RAPASSFYQKENKPYSSNSPAGLQKNPIWHEGENLYGMQITRLNSPEREHLAFPDLPYSPYAVQKGAHLFHLQRGPCAGGSTGPKDNPLALQDYTP	418
bottlenose_dolphin	KAPASSFYQKENKPYSSNSPAGLQKNPIWHEGENLYGMQITRLNSPEREHLAFPDLPYSPYAVQKGAHLFHLQRGPCAGGSTGPKDNPLALQDYTP	873
finless_porpoise	KVPASSFYQKENKPYSSNSPAGLQKNPIWHEGENLYGMQITRLNSPETEHLVFPDLPYSPYAVQKGAHLFHLQRGPCAGGSTGPKDNPLALQDYTP	873
cow	KAPASSFYQKENKPYSSNSPAGLQKNPIWHEGENLYGMQITRLNSPEGGHLAFPDLPYSPYAVQKGAHLFHLQRGPCAGGSTGPKDNPLALQDYTP	894
pig	KTPSSFYQKENKPYSSNSPAGLQKNPTWHEGENLYGMQITRLNSPERDHLAFPDLPYSPYAVQKGAHLFHLQRGPCAGGSTGPKDNPLALQDYTP	888
minke_whale	SFGLAPGENQDTSPIVYEDSHTKHARHTISPTSLPQQRNSSEKRLPGESQNSPFRDDVSL*	501
fin_whale	SFGLAPGENQDTSPIVYEDSHTKHARHTISPTSLPQQRNSSEKRLPGESQNSPFRDDVSL*	500
bottlenose_dolphin	SFGLALGENQDTSPIVYEDSQTTHVRHTISPTSLAQQRNSSEKRLPGESQNSPFRDDMSTLARNTPRSIN--QLGQRG	972
finless_porpoise	SFGLAPGENQDTSPIVYEDSQTTHVRHTISPTSLAQQRNSSEKRLPGESQNSPFRDDMSTLARNTPRSIN--QLGQRG	972
cow	FFGLAPGENQDTSPIVYEDSHTKHARHTISPTSLPQQRNSSEKRLPGESQNSPFRDDVSTLARNTPCSIN--QLSQRG	994
pig	SFGLAPGENDTSPIVYEDSHTKHARHTISPTSLPQQRNSSEKRLPGESQNSPFRDDVSTLARNTPCSVKSQLSQRG	988
minke_whale	LGGEEN--ILEQIFEGNPLNERTVDLTPEQLVIGIPDEGPKPEGIPSEVQNGEGERQ--RATSLQLPCFGSKLTKHHTSSGTTPSSFGRQGSFDGDPIMP	599
fin_whale	LGGEEN--ILEQIFEGNPLNERTVDLTPEQLVIGIPDEGPKPEGIPSEVQNGEGERQ--RATSLQLPCFGSKLTKYHTSSGTTPSSFGRQGSFDGDPIMP	598
bottlenose_dolphin	LGGEEN--ILEQIFEGNQLNERTVDLTPEQLVIGIPDEGPKPEGIPSEVQNGEGERQ--RPSSILQLPCFGSKLTKYHTSSGTTPSSIGRQGSFDGDPIMP	1070
finless_porpoise	LGGEEN--ILEQIFEGNQLNERTVDLTPEQLVIGIPDEGPKPEGIPSEVQNGEGERQ--RPSSILQLPCFASKLTKYHTSSGTTPSSIGRQGSFDGDPIMP	1070
cow	L--EDGNHVLEQTLNQLNERTVDLTPEQLVMDTPDEGPKPEGIPSEVQNGGKRQQRSTILKLPDFSKLTKYHTSSGTTPSSIGRQGSFDGDPIMP	1093
pig	LGGDGNVLEQIFEGNQLNERTVDLTPEQLVGTDEEPRPEGIPNEMQNGEGERQ--RATSLQLPCFGSKLTKYHTSSGTTPSSIGRQGSFDGDPIMP	1088

**Supplementary Figure 34 | Alignments of *ENAM* genes.** The *ENAM* gene of the minke and fin whales had two and three early stop codons (red arrows), respectively. The bottlenose dolphin and finless porpoise did not have early stop codons.

First early stop codon in minke whale *ENAM* gene



First early stop codon in fin whale *ENAM* gene



Non stop codon in minke whale *ENAM* gene



Second early stop codon in fin whale *ENAM* gene



Second early stop codon in minke whale *ENAM* gene



Third early stop codon in fin whale *ENAM* gene



**Supplementary Figure 35 | Sanger sequencing results for the *ENAM* gene in the baleen whale. Red boxes are early stop codons in the baleen whales.**



```

minke_whale      MGTWILFACLLGAAFSMPLPPHPGHPGYINFRDE--YFQVVRIDGTASVLTPLKWFQNI 57
fin_whale       MGTWILFACLLGAAFSMPLPPHPGHPGYINFSYE--HFQVVRIDGTASVLTPLKWFQNI 57
bottlenose_dolphin MGTWILFACLLGAAFSMPLPPHPGHPGYINFSYE--YFQVVRIDGTASVLTPLKWFQNI 58
finless_porpoise MGTWILFACLLGAAFSMPLPAHPGHPGYINFSYE--YFQVVRIDGTASVLTPLKWFQNI 58
pig            MGTWILFACLLGAAFSMPLPPHPGHPGYINFSYE-----VLTPLKWFQNI 46
mouse          MGTWILFACLLGAAFAMPLPPHPGSPGYINLSYEKSHSQAINTDRALVLTPLKWFQSMI 60

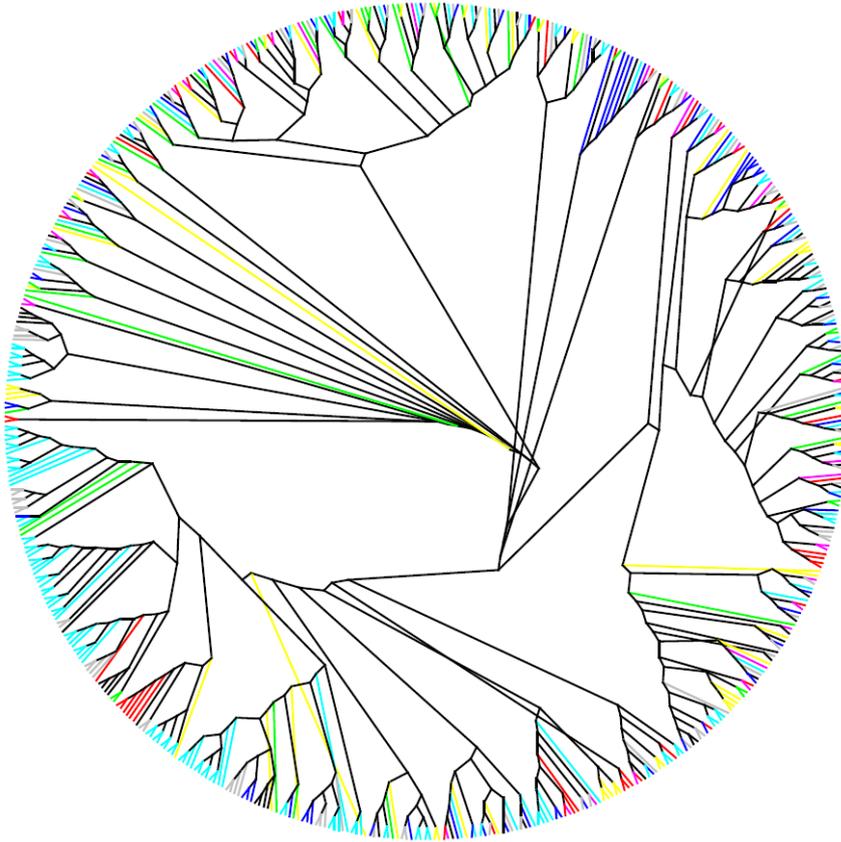
minke_whale      RHP-PSYGYEPMGGRLHHRIIPVVSQQH---ALPPHHHVMEPAN----- 98
fin_whale       RHP-PSYGYEPMGGWLHHQIIPVVSQQP---ALPPHHHFMVPAN----- 98
bottlenose_dolphin RHPYPSYGYEPMGGWLHHQIIPVVSQQA---ALQPHHHFPMVPAQQPVGPGQTMMP 114
finless_porpoise RHPYPSYGYEPMGGWLHHQIIPVVSQQA---ALQPHHHFPMVPAQQPVVPGQTMMP 114
pig            RHPYTSYGYEPMGGWLHHQIIPVVSQQTPQSHALQPHHHIPMVPAQQPGIPQQPMPL 106
mouse          RQPYPSYGYEPMGGWLHHQIIPVLSQQHPPSHTLQPHHHLPVVPAAQPVAPQQPM 120

minke_whale      ----- 153
fin_whale       ----- 153
bottlenose_dolphin QHSMTPSQPHQPHLP---VPAQQPVQPQPHPP-----LLQPPLPPM 153
finless_porpoise QHSMTPSQPHQPHLP---VPAQQPVQPQPHPP-----LLQPPLPPM 153
pig            QHSMTPTQHHQPNLP----LPAQQPFQPQVQPQPHQPLQPQSPMHPIQPLLPQPPL 162
mouse          HHSMTPTQHHQPNIPPSAQQPFQPQVQPQIIPQSHQPMQPSLHPMQPLAQQPPL 180

minke_whale      ----- 183
fin_whale       ----- 183
bottlenose_dolphin FPMQPLPPMLPDLPLEAWPATDKTKREEVD 183
finless_porpoise FPVQPLPPMLPDLPLEAWPATDKTKREEVD 183
pig            FSMQSL---LPDLPLEAWPATDKTKREEVD 189
mouse          FSMQPLSPILPELPLEAWPATDKTKREEVD 210

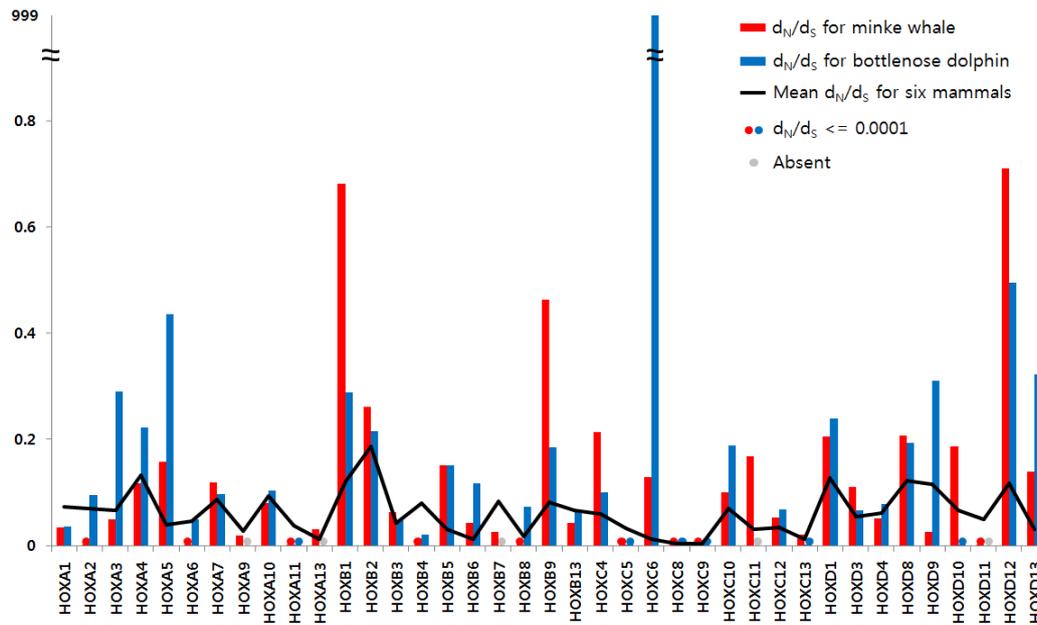
```

**Supplementary Figure 37 | Alignments of *AMEL* genes.** The *AMEL* gene of the minke and fin whales had one stop codon (red arrow). The bottlenose dolphin and finless porpoise did not have early stop codons.



**Supplementary Figure 38 | Contracted keratin-related gene families in the whale lineage.**

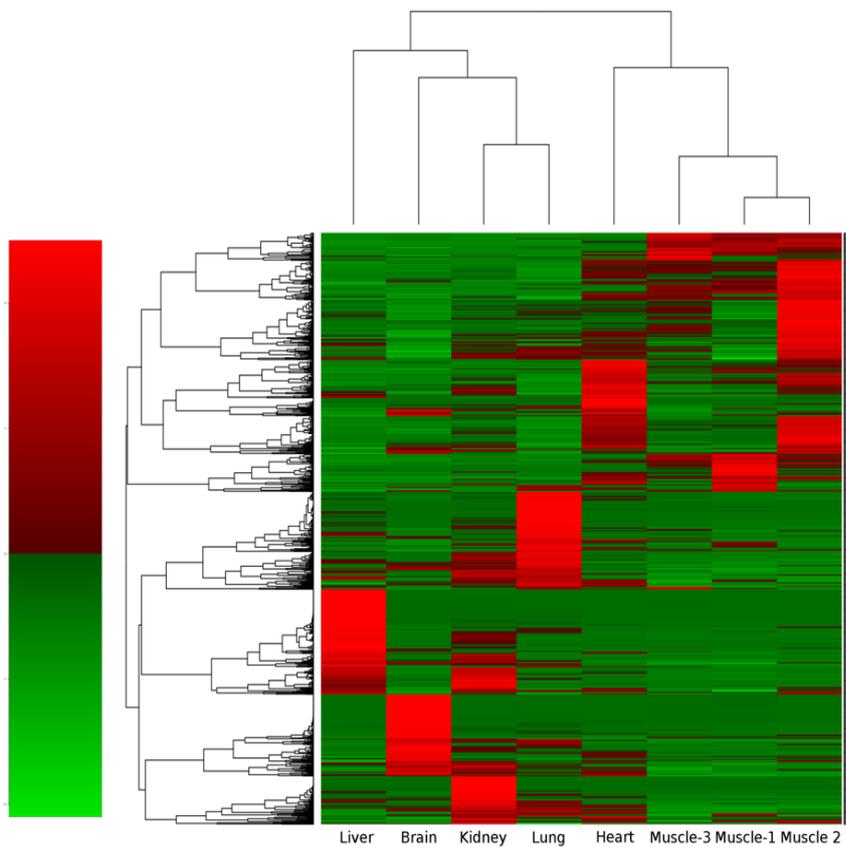
Compared with other mammalian species, the minke whale and bottlenose dolphin had evolutionarily contracted keratin-related genes (*KRT* and *KRTAP* gene families). The red, magenta, silver, black, blue, cyan, green, and yellow nodes are keratin-related genes in the minke whale, bottlenose dolphin, cow, pig, dog, human, opossum, and platypus, respectively. A total of 573 genes in 52 gene families among eight mammalian species were used in this analysis. The numbers of keratin-related genes in each species are 49 (minke whale), 33 (bottlenose dolphin), 89 (cow), 60 (pig), 75 (dog), 163 (human), 47 (opossum), and 57 (platypus).



**Supplementary Figure 39 |  $d_N/d_S$  ratios for the *Hox* genes.** The  $d_N/d_S$  for the *Hox* genes in the minke whale and bottlenose dolphin were from the free ratio (M1) model by using cow, pig, human, and mouse as background branches. The mean  $d_N/d_S$  ratios for the *Hox* genes in the six mammals (minke whale, bottlenose dolphin, cow, pig, human, and mouse) were from the one ratio (M0) model.

minke_whale	MDHAEENEILAAQRYYVERPFSHPVLERLHKHKDSISDGDKLQAFCTPKKIRNIIYMFLPIKWLPAFRFKEYVLDLVSIGSTGVLQPLQGLAFAMLAAPPVFLYSSFYVPV	120
fin_whale	MDHAEENEILAAQRYYVERPFSHPVLERLHKHKDSISDGDKLQAFCTPKKIRNIIYMFLPIKWLPAFRFKEYVLDLVSIGSTGVLQPLQGLAFAMLAAPPVFLYSSFYVPV	120
bottlenose_dolphin	MDHAEENEILAAQRYYVERPFSHPVLERLHKHKDSISDGDKLQAFCTPKKIRNIIYMFLPIKWLPAFRFKEYVLDLVSIGSTGVLQPLQGLAFAMLAAPPVFLYSSFYVPV	120
finless_porpoise	MDHAEENEILAAQRYYVERPFSHPVLERLHKHKDSISDGDKLQAFCTPKKIRNIIYMFLPIKWLPAFRFKEYVLDLVSIGSTGVLQPLQGLAFAMLAAPPVFLYSSFYVPV	120
cow	MDHAEENEILAAQRYYVERPFSHPVLERLHKHKDSISDGDKLQAFCTPKKIRNIIYMFLPIKWLPAFRFKEYVLDLVSIGSTGVLQPLQGLAFAMLAAPPVFLYSSFYVPV	120
pig	MDHAEENEILAAQRYYVERPFSHPVLERLHKHKDSISDGDKLQAFCTPKKIRNIIYMFLPIKWLPAFRFKEYVLDLVSIGSTGVLQPLQGLAFAMLAAPPVFLYSSFYVPV	120
human	MDHAEENEILAAQRYYVERPFSHPVLERLHKHKDSISDGDKLQAFCTPKKIRNIIYMFLPIKWLPAFRFKEYVLDLVSIGSTGVLQPLQGLAFAMLAAPPVFLYSSFYVPV	120
minke_whale	IMYCFPGTSRHSIGPFAVISMIGGAVRLVPDDIVIPGGVATNGTEARDALRVKVMVSTLLSGIIQ-----VRFQFVAIYLTEPLVRGFTTAAAVHVFTSMLKYLFGVKTKRYSGI	235
fin_whale	IMYCFPGTSRHSIGPFAVISMIGGAVRLVPDDIVIPGGVATNGTEARDALRVKVMVSTLLSGIIQ-----VRFQFVAIYLTEPLVRGFTTAAAVHVFTSMLKYLFGVKTKRYSGI	235
bottlenose_dolphin	IMYCFPGTSRHSIGPFAVISMIGGAVRLVPDDIVIPGGVATNGTEARDALRVKVMVSTLLSGIIQ-----VRFQFVAIYLTEPLVRGFTTAAAVHVFTSMLKYLFGVKTKRYSGI	190
finless_porpoise	IMYCFPGTSRHSIGPFAVISMIGGAVRLVPDDIVIPGGVATNGTEARDALRVKVMVSTLLSGIIQ-----VRFQFVAIYLTEPLVRGFTTAAAVHVFTSMLKYLFGVKTKRYSGI	190
cow	IMYCFPGTSRHSIGPFAVISMIGGAVRLVPDDIVIPGGVATNGTEARDALRVKVMVSTLLSGIIQCLGVCRFQFVAIYLTEPLVRGFTTAAAVHVFTSMLKYLFGVKTKRYSGI	240
pig	IMYCFPGTSRHSIGPFAVISMIGGAVRLVPDDIVIPGGVATNGTEARDALRVKVMVSTLLSGIIQCLGVCRFQFVAIYLTEPLVRGFTTAAAVHVFTSMLKYLFGVKTKRYSGI	240
human	IMYCFPGTSRHSIGPFAVISMIGGAVRLVPDDIVIPGGVATNGTEARDALRVKVMVSTLLSGIIQCLGVCRFQFVAIYLTEPLVRGFTTAAAVHVFTSMLKYLFGVKTKRYSGI	240
minke_whale	FSVYVYMTVLQSTVAVLQNKLNKLVCSLGVGLMVFGLLLGGKEFNERFKEKLPAPIPLEFFAVVMGTGISAGFNLHESYNDVVGTLPLG-----LFHLVYDAIAIAIVGFSV	345
fin_whale	FSVYVYMTVLQSTVAVLQNKLNKLVCSLGVGLMVFGLLLGGKEFNERFKEKLPAPIPLEFFAVVMGTGISAGFNLHESYNDVVGTLPLG-----LFHLVYDAIAIAIVGFSV	345
bottlenose_dolphin	-----STVAVLQNKLNKLVCSLGVGLMVFGLLLGGKEFNERFKEKLPAPIPLEFFAVVMGTGISAGFNLHESYNDVVGTLPLG-----LFHLVYDAIAIAIVGFSV	289
finless_porpoise	-----STVAVLQNKLNKLVCSLGVGLMVFGLLLGGKEFNERFKEKLPAPIPLEFFAVVMGTGISAGFNLHESYNDVVGTLPLG-----LFHLVYDAIAIAIVGFSV	289
cow	FSVYVY-----STVAVLQNKLNKLVCSLGVGLMVFGLLLGGKEFNERFKEKLPAPIPLEFFAVVMGTGISAGFNLHESYNDVVGTLPLG-----LFHLVYDAIAIAIVGFSV	353
pig	FSVYVY-----STVAVLQNKLNKLVCSLGVGLMVFGLLLGGKEFNERFKEKLPAPIPLEFFAVVMGTGISAGFNLHESYNDVVGTLPLG-----LFHLVYDAIAIAIVGFSV	353
human	FSVYVY-----STVAVLQNKLNKLVCSLGVGLMVFGLLLGGKEFNERFKEKLPAPIPLEFFAVVMGTGISAGFNLHESYNDVVGTLPLG-----LFHLVYDAIAIAIVGFSV	353
minke_whale	TISMAKTLANKHGQVQDGNQELIALGLCNSIGSLFQTFISCSLSRSLVQEGTGKTLQAGCLASMLILLVILATGFLFESLPQAVLSAIVIVNLKGMFMQFSDLPFFWRTSKIELTIWL	465
fin_whale	TISMAKTLANKHGQVQDGNQELIALGLCNSIGSLFQTFISCSLSRSLVQEGTGKTLQAGCLASMLILLVILATGFLFESLPQAVLSAIVIVNLKGMFMQFSDLPFFWRTSKIELTIWL	465
bottlenose_dolphin	-----KTAHYSNGNQELIALGLCNSIGSLFQTFISCSLSRSLVQEGTGKTLQAGCLASMLILLVILATGFLFESLPQAVLSAIVIVNLKGMFMQFSDLPFFWRTSKIELTIWL	399
finless_porpoise	-----KTAHYSNGNQELIALGLCNSIGSLFQTFISCSLSRSLVQEGTGKTLQAGCLASMLILLVILATGFLFESLPQAVLSAIVIVNLKGMFMQFSDLPFFWRTSKIELTIWL	399
cow	TISMAKTLANKHGQVQDGNQELIALGLCNSIGSLFQTFISCSLSRSLVQEGTGKTLQAGCLASMLILLVILATGFLFESLPQAVLSAIVIVNLKGMFMQFSDLPFFWRTSKIELTIWL	473
pig	TISMAKTLANKHGQVQDGNQELIALGLCNSIGSLFQTFISCSLSRSLVQEGTGKTLQAGCLASMLILLVILATGFLFESLPQAVLSAIVIVNLKGMFMQFSDLPFFWRTSKIELTIWL	473
human	TISMAKTLANKHGQVQDGNQELIALGLCNSIGSLFQTFISCSLSRSLVQEGTGKTLQAGCLASMLILLVILATGFLFESLPQAVLSAIVIVNLKGMFMQFSDLPFFWRTSKIELTIWL	473
minke_whale	TFVSSSLFLGLDYLITAVIIALLTVIYRTQSPSYKVLGQLPDDTVYIDIDAYEEVKEIPGKIFQINAPIYYANSOLYSSALKRKTGVNPAVIMGARRKAMKRYAKEVGNANMANATV	585
fin_whale	TFVSSSLFLGLDYLITAVIIALLTVIYRTQSPSYKVLGQLPDDTVYIDIDAYEEVKEIPGKIFQINAPIYYANSOLYSSALKRKTGVNPAVIMGARRKAMKRYAKEVGNANMANATV	585
bottlenose_dolphin	TFVSSSLFLGLDYLITAVIIALLTVIYRTQSPSYKVLGQLPDDTVYIDIDAYEEVKEIPGKIFQINAPIYYANSOLYSSALKRKTGVNPAVIMGARRKAMKRYAKEVGNANMANATV	519
finless_porpoise	TFVSSSLFLGLDYLITAVIIALLTVIYRTQSPSYKVLGQLPDDTVYIDIDAYEEVKEIPGKIFQINAPIYYANSOLYSSALKRKTGVNPAVIMGARRKAMKRYAKEVGNANMANATV	519
cow	TFVSSSLFLGLDYLITAVIIALLTVIYRTQSPSYKVLGQLPDDTVYIDIDAYEEVKEIPGKIFQINAPIYYANSOLYSSALKRKTGVNPAVIMGARRKAMKRYAKEVGNANMANATV	593
pig	TFVSSSLFLGLDYLITAVIIALLTVIYRTQSPSYKVLGQLPDDTVYIDIDAYEEVKEIPGKIFQINAPIYYANSOLYSSALKRKTGVNPAVIMGARRKAMKRYAKEVGNANMANATV	593
human	TFVSSSLFLGLDYLITAVIIALLTVIYRTQSPSYKVLGQLPDDTVYIDIDAYEEVKEIPGKIFQINAPIYYANSOLYSSALKRKTGVNPAVIMGARRKAMKRYAKEVGNANMANATV	593
minke_whale	KV-DAEVDGEDGTPKEEENEIKYPPVIVKSTLPEELQRMPMPGDNVHTIILDFQTQVNFIDSVGVKTLAGIVKEYGDVGIYVYLAGCSAQVNDLSTRNRFENPALLELLFHSIHDVAVLG	704
fin_whale	KV-DAEVDGEDGTPKEEENEIKYPPVIVKSTLPEELQRMPMPGDNVHTIILDFQTQVNFIDSVGVKTLAGIVKEYGDVGIYVYLAGCSAQVNDLSTRNRFENPALLELLFHSIHDVAVLG	704
bottlenose_dolphin	KV-DAEVDGEDGTPKEEENEIKYPPVIVKSTLPEELQRMPMPGDNVHTIILDFQTQVNFIDSVGVKTLAGIVKEYGDVGIYVYLAGCSAQVNDLSTRNRFENPALLELLFHSIHDVAVLG	639
finless_porpoise	KV-DAEVDGEDGTPKEEENEIKYPPVIVKSTLPEELQRMPMPGDNVHTIILDFQTQVNFIDSVGVKTLAGIVKEYGDVGIYVYLAGCSAQVNDLSTRNRFENPALLELLFHSIHDVAVLG	639
cow	KV-DAEVDGEDGTPKEEENEIKYPPVIVKSTLPEELQRMPMPGDNVHTIILDFQTQVNFIDSVGVKTLAGIVKEYGDVGIYVYLAGCSAQVNDLSTRNRFENPALLELLFHSIHDVAVLG	712
pig	KV-DAEVDGEDGTPKEEENEIKYPPVIVKSTLPEELQRMPMPGDNVHTIILDFQTQVNFIDSVGVKTLAGIVKEYGDVGIYVYLAGCSAQVNDLSTRNRFENPALLELLFHSIHDVAVLG	712
human	KA-DAEVDGEDGTPKEEENEIKYPPVIVKSTLPEELQRMPMPGDNVHTIILDFQTQVNFIDSVGVKTLAGIVKEYGDVGIYVYLAGCSAQVNDLSTRNRFENPALLELLFHSIHDVAVLG	712

**Supplementary Figure 40 | Alignments of prestin genes.** The bottlenose dolphin and finless porpoise sequences showed the highest frequency of amino acid changes.



**Supplementary Figure 41 | Heat map of the genes enriched in each organ.** The heat map was generated by hierarchical clustering and complete linkage of the top 500 most highly expressed genes. Distances were calculated using Pearson's correlation coefficient. Genes that were highly expressed in an organ-specific manner are shown in red.

## Supplementary Tables

**Supplementary Table 1 | Sample information.** Four minke whales, a fin whale, a bottlenose dolphin, and a finless porpoise were used in this study. A male minke whale was used to assemble the draft genome.

Sample	Sex	Origin
Minke whale	Male	South Korea
Minke whale01	Female	South Korea
Minke whale02	Female	South Korea
Minke whale03	Female	South Korea
Fin whale	Female	U.S. West coast
Bottlenose dolphin	Male	South Korea
Finless porpoise	Unknown	South Korea

**Supplementary Table 2 | Raw read filtering statistics.** The raw reads of six cetaceans were generated by Illumina Hiseq2000.

Sample	The number of raw read pairs	The number of proper read pairs	The percentage of proper read pairs	Estimated sequencing depth from raw read pairs	Estimated sequencing depth from proper read pairs
Minke whale01	544,765,050	516,266,141	94.77	39.43	37.37
Minke whale02	512,157,463	496,051,180	96.86	37.07	35.90
Minke whale03	438,408,187	413,853,352	94.40	31.73	29.95
Fin whale	619,941,025	578,742,088	93.35	44.87	41.89
Bottlenose dolphin	598,122,699	580,006,136	96.97	43.29	41.98
Finless porpoise	553,390,610	523,612,238	94.62	40.05	37.90

**Supplementary Table 3 | Details of the seven libraries used by the minke whale genome sequencing project.** A total of 354.56 Gb of short reads were generated from 14 libraries (18 lanes) to assemble the minke whale genome. Insert sizes include paired-end read lengths.

Paired-end libraries	Insert size	Read length (bp)	Total data (Gb)	Sequence Depth (X)	Physical Depth (X)
Illumina	170bp	100	78.61	28.48	24.21
Reads	500bp	100	27.80	10.07	25.18
	800bp	100	81.30	29.46	117.83
	2Kb	49	56.29	20.39	416.22
	5Kb	49	40.57	14.70	749.96
	10Kb	49	38.01	13.77	1405.28
	20Kb	49	31.98	11.59	2364.69
Total	-----	-----	354.56	128.46	5103.37

**Supplementary Table 4 | Details of the filtered sequencing data for the minke whale genome *de novo* assembly.** In order to increase the assembly quality, the sequences were processed by filtering steps. Insert sizes include paired-end read lengths.

Paired-end libraries	Insert size	Read length	Total data (Gb)	Sequence Depth (X)	Physical Depth (X)
Illumina Reads	170bp	100	70.46	25.53	21.70
	500bp	100	24.61	8.92	22.29
	800bp	100	65.56	23.75	95.01
	2Kb	49	41.22	14.93	304.79
	5Kb	49	23.67	8.58	437.56
	10Kb	49	20.55	7.45	759.76
	20Kb	49	7.33	2.66	542.00
Total	-----	-----	253.40	91.82	2183.11

**Supplementary Table 5 | Library mapping statistics for the assembled minke whale genome.** Each library was mapped to the assembled minke whale genome. SD is standard deviation. Insert sizes include paired-end read lengths.

Library Insert size	Observed insert size			The number of paired-ends where both ends uniquely mapped	Genome coverage (%)
	Mode	Mean	SD		
170bp	165	165	9	242,404,598	93
500bp	479	478	12	72,763,633	92
800bp	786	786	16	220,210,187	92
2Kb	2,327	2,300	425	116,152,608	93
5Kb	5,377	5,027	1,699	67,393,943	92
10Kb	10,897	10,874	1,676	77,665,397	84
20Kb	19,658	19,611	3,778	17,032,143	69

**Supplementary Table 6 | Details of the 17-mer statistical analysis.** Peak\_depth means the depth value at the peak of the distribution curve of depth-frequency.

K-mer	K-mer_num	Peak_depth	Genome_size	Used_base	Used_read	Depth
17	69,083,053,017	25	2,763,322,121	86,114,951,980	1,013,074,370	30.808

**Supplementary Table 7 | Data after error correction.** Randomly occurred sequencing errors were corrected using K-mer frequency information.

Pair-end libraries	Insert Size	Total Data (Gb)	Read Length (bp)	Sequence Coverage (X)	Physical Coverage (X)
Illumina	170bp	69.14	98.74	25.05	21.29
Reads	500bp	23.05	96.53	8.35	20.88
	800bp	64.41	98.82	23.34	93.35
Total	---	156.6	---	56.74	135.52

**Supplementary Table 8 | Details of the assembled sequence length.** Contig represents the final contig after filling the gap in the intra-scaffold. The above table includes length information for the assembled genome sequence such as the scaffold and final contig's N50 to N90, the total length, and maximum length. The scaffolds with a length of < 100bp were not included in the statistics.

	Contig		Scaffold	
	Size (bp)	Number	Size (bp)	Number
N90	6,538	102,280	2,659,324	208
N80	10,570	75,033	5,028,618	143
N70	14,395	56,477	7,819,582	105
N60	18,298	42,316	10,648,577	78
N50	22,571	31,010	12,843,668	57
Longest	178,666	-	51,448,160	-
Total Size	2,297,917,021	-	2,442,893,294	-
Total Number (>100bp)	-	278,792	-	104,325
Total Number (>2kb)	-	147,743	-	2,310

**Supplementary Table 9 | DNA read mapping statistics for the assembled genome.** The library with insert-size of 800bp was mapped to the assembled minke whale genome.

Sample	The Number of All Reads	The Number of Mapped Reads	The Number of Unmapped Reads	The Number of Non-redundant Mapped Reads	The Percentage of Mapped Reads	The Percentage of Unmapped Reads	The Percentage of Non-redundant Mapped Reads	Mapping Depth
Minke whale	771,736,654	702,292,494	69,444,160	600,900,800	91.00	9.00	85.56	26

**Supplementary Table 10 | RNA read mapping statistics for the assembled genome**

Tissue	Sequencing			Selected			Uniquely Mapped
	Raw	Filtered	Selected	Total	Mapped	Unmapped	
Lung	30,156,598	481,174 1.60%	29,675,424 98.40%	29,675,424	25,571,049 86.20%	4,104,375 13.80%	24,261,842 81.80%
Brain	51,470,260	891,446 1.70%	50,578,814 98.30%	50,578,814	42,632,914 84.30%	7,945,900 15.70%	41,007,087 81.10%
Heart	86,122,156	1,257,094 1.50%	84,865,062 98.50%	84,865,062	52,348,831 61.70%	32,516,231 38.30%	49,814,329 81.80%
Kidney	57,581,784	858,092 1.50%	56,723,692 98.50%	56,723,692	39,050,026 68.80%	17,673,666 31.20%	37,508,046 66.10%
Liver	47,846,388	754,988 1.60%	47,091,400 98.40%	47,091,400	40,229,415 85.40%	6,861,985 14.60%	38,605,822 81.80%
Muscle01	58,644,336	985,452 1.70%	57,658,884 98.30%	57,658,884	48,402,577 83.90%	9,256,307 16.10%	46,464,922 80.60%
Muscle02	95,161,644	1,617,084 1.70%	93,544,560 98.30%	93,544,560	71,653,215 76.60%	21,901,345 23.40%	68,635,380 81.80%
Muscle03	31,034,064	559,748 1.80%	30,474,316 98.20%	30,474,316	22,784,253 74.80%	7,690,063 25.20%	21,734,477 71.30%

**Supplementary Table 11 | Mapping statistics for unmapped RNA reads of the minke whale transcriptome.** The unmapped reads\* mean RNA reads that were not mapped to the minke whale genome.

Tissue	Reference genome	Mapping statistics for unmapped minke whale transcriptome reads				
		Mapped reads		Unmapped reads		The number of total unmapped reads*
Lung	Cow	136,699	3.3%	3,967,676	96.7%	4,104,375
	Pig	130,122	3.2%	3,974,253	96.8%	
	Human	106,769	2.6%	3,997,606	97.4%	
Brain	Cow	314,566	4.0%	7,631,334	96.0%	7,945,900
	Pig	298,120	3.8%	7,647,780	96.2%	
	Human	255,922	3.2%	7,689,978	96.8%	
Heart	Cow	617,181	1.9%	31,899,050	98.1%	32,516,231
	Pig	599,230	1.8%	31,917,001	98.2%	
	Human	484,676	1.5%	32,031,555	98.5%	
Kidney	Cow	241,415	1.4%	17,432,251	98.6%	17,673,666
	Pig	227,190	1.3%	17,446,476	98.7%	
	Human	198,088	1.1%	17,475,578	98.9%	
Liver	Cow	300,322	4.4%	6,561,663	95.6%	6,861,985
	Pig	284,646	4.1%	6,577,339	95.9%	
	Human	269,084	3.9%	6,592,901	96.1%	
Muscle01	Cow	490,620	5.3%	8,765,687	94.7%	9,256,307
	Pig	443,862	4.8%	8,812,445	95.2%	
	Human	529,937	5.7%	8,726,370	94.3%	
Muscle02	Cow	607,297	2.8%	21,294,048	97.2%	21,901,345
	Pig	586,462	2.7%	21,314,883	97.3%	
	Human	488,947	2.2%	21,412,398	97.8%	
Muscle03	Cow	689,916	9.0%	7,000,147	91.0%	7,690,063
	Pig	665,062	8.6%	7,025,001	91.4%	
	Human	594,737	7.7%	7,095,326	92.3%	

**Supplementary Table 12 | Assessment of gene coverage by assembled minke whale transcripts**

Dataset	Number	Total Length (bp)	Covered by Assembly (%)	With >90% Sequence in one Scaffold		With >50% Sequence in one Scaffold	
				Number	Percent (%)	Number	Percent (%)
All	537,572	756,849,198	96.73	483,789	88.99	531,033	98.78
>200bp	537,572	756,849,198	96.73	483,789	88.99	531,033	98.78
>500bp	245,930	669,427,722	96.64	215,723	87.71	243,447	98.99
>1000bp	163,720	612,274,264	96.57	141,209	86.25	162,043	98.97

**Supplementary Table 13 | Evaluation of the completeness of minke whale genome assembly using core mammalian gene mapping approach (CEGMA).** Above 98.6% of conserved genes were found in the assembled scaffolds. A KOG is a mammalian orthologous gene. There is a total of 458 mammalian KOGs.

Parameter	KOG number in minke whale	Percent (%)
Total KOGs matched	452	98.69
KOG aligned to one gene	416	90.83
KOG aligned to several genes	13	2.84
KOG aligned to no gene	23	5.02

**Supplementary Table 14 | SNVs and small indels statistics for the assembled whale genome.** SNVs and indels were found by mapping the library with insert-size of 800bp to the assembled minke whale genome.

Sample	The Number of All Variants	The Number of Homozygous SNV Locus	The Number of Heterozygous SNV Locus	The Number of Small Indels	Mapping Depth
Minke whale	2,168,536	642,732	1,293,784	232,020	26

**Supplementary Table 15 | Heterozygous SNV validation using Sanger sequencing.** Randomly selected heterozygous SNV positions were validated by Sanger sequencing method, and 98.6% (139) were true heterozygous SNVs.

scaffold	position	Ref. Allele	Hetero-zygous Allele	Forward Primer	Reverse Primer	SNV status
scaffold63	950577	T	G	GCTTTCCAGAAGACACTCGTAGC	TCCAGACTAAGCGCATCTTGAG	valid
scaffold99	3450916	G	A	ACCCAGCATATTCAGGGGATATT	GGATTGCAITTCAGTTTGTCT	valid
scaffold48	3003253	T	C	GGACTTACAGGCTCCTCAGTTGA	CAAACTAGCATTGATTCCTTG	valid
scaffold99	5562858	G	T	CACACTGAGGTGAGTCCCTCCT	TGCTGTCTCCATGACACATCT	valid
scaffold18	4340888	G	C	AGCCCCAACTATGCTATTCAACA	CGCTGTTTTAGGGTCCCTTCTAT	valid
scaffold48	5274806	A	G	GCTGACATAGGTGTTACAATGG	TAGATTGTGAAGTTTCCCCCA	valid
scaffold63	5637364	G	T	TTGCCGTTGCTTTTCCCTTTT	GTTTGCTAATCTTTGGCCCTC	valid
scaffold52	10796204	G	C	TCGACACTGTAAAAGCTCCACAA	ATCATCCGCAGAAGTTTGTCTC	valid
scaffold117	2914847	C	G	GTGACCTTCCAAACCACATCTTC	ACGATGACATGATTGATGTTTGC	valid
scaffold129	5460427	T	C	AGAATTTGGATGCTATTGCCGTA	CTGCTCAGTGCCTAATGGTTAC	valid
scaffold135	326015	A	G	ATTGTGGCAGTTGTGCTGTAGAA	CCCGTGTTTATGAATAACTTGGC	valid
scaffold129	4119001	C	T	TGATCCGTTACATAATGACACC	GGCCTATCCTTCAAACCTTGGTT	valid
scaffold41	6377102	C	T	TAGTAGTATGCGAAAAGGGCTCA	ATTGGCATGTGCTCACTCAATAA	valid
scaffold99	5934329	T	A	CCCTTACGTGATCAGACCAATAGG	TTGCCCTCAATAGATACCAAGT	valid
scaffold23	1276124	A	G	TTTCAGTTGACAAATGGGCTGATA	GGGAATTATTGGGAAAAGTGAGC	valid
scaffold23	878343	G	A	CCTCATGAGAACCCTCAGAAAGAA	GGCTTTTGCTAAGCTGTGCTCA	valid
scaffold101	620483	T	C	AAATCCCTCCTCCTGAAGTTTTC	ACCTACTAAAGGCCACGTGTCA	valid
scaffold124	1637179	T	C	GGTACACAGGATGATGGAAGTGAG	TTTCTCAAATTTGGGACAGAAA	valid
scaffold47	8560199	A	G	TTCTTCGAGAATCCACCAGAAAG	AAAAGCACCTAAGATTCAAGCCC	valid
scaffold81	9305528	C	T	CCTCTTTGCCCTTCTAGTCCCAT	ATGGAACAATTTGGTGTGCTCT	valid
scaffold13	4368086	C	T	ATGTGCAAAACTGTCCATTCTCT	ATGCCACAAAATAAAATGTTGCC	valid
scaffold180	2378336	G	T	TTGCCAAGTTGACCACTAAGGAT	TGCCTAACATCACCAGTGAAGAA	valid
scaffold88	3507369	A	G	TGCTCCTGACGTATTGAATCAGA	GAATCTCAATCTTCCAGCTGAA	valid
scaffold116	4348127	G	A	GCTCAGAACCCGTTCTACTCTTTA	GCTGACTACTCTCTCCATTGCC	valid
scaffold1	41965143	T	C	CTGGTTTAAAGTTGCAATTTCTGG	GCACGATAACTCAAGACTCAAAGA	valid
scaffold1	9148260	G	T	TTCTGACCGTGATGCTCAACTTA	TTATAAACAGGGAGCAAGCAAGC	valid
scaffold76	5750308	G	C	GGCTTAAAAGAAAAACAGCCGTT	TGGGCATAATTTCTTTTATGCG	valid
scaffold17	1489718	T	C	GGTCCCTTACGTGGGTTCTTAAC	GAAGTCCCCAAAAGGAAACATCT	valid
scaffold35	15787752	A	G	TGTAGTCTCATTCTCTCCCAA	TTTCTCCACCTTCAAAGTCAACC	valid
scaffold17	10419286	G	A	CAACCTCTCTTGTCTCTGGA	CCCTTGCCTTCTGTATAAAT	valid
scaffold57	2951272	A	G	TGCTGGAGATACAACCTTGACCA	AGTCTGTCTACTCGCCGATATG	valid
scaffold72	7228756	T	A	TCFTTCCAGTGTGAAACATTA	TCCTTTGAGAAGTGAAGGCTCG	valid
scaffold57	11179089	C	G	GTGGTGTAGCCAAGTTGTAAGG	CAGGGCTTTTGTGTTTCTCAAC	valid
scaffold1	1765592	A	C	AACTCTGAGCCTGTCTATTTCCAG	TGCATGCTCTCTGGGTTTATTT	valid
scaffold54	10448272	T	C	CCCTCATTTAGTACTGTCCAGA	TTTTGAGCAGAGAATTTCAGGAGG	valid
scaffold75	4644550	G	A	TCCTAGGTTTTTCTTGATGGAG	GTGCAGCTCAAACAAAACAATG	valid
scaffold1	28472046	T	A	TCATCATCTCTCTGATTGTCT	TTCTGTCTCGTGCCTTACTTC	valid
scaffold43	13227237	A	T	AGATTTGTACGCTTCTTAGGGGG	ACGATCTTGGCAACATTTTGAAT	valid
scaffold30	17450679	A	G	CCTCTTACCAGCTGTGTCTCTC	TTTTGGTTGTTTTGGAAGTGCT	valid
scaffold30	2704557	A	C	TGGTAAGGTCATGGTTGACACTG	CTAATCACTCCTCAAAGGCTCCA	valid
scaffold12	11209259	G	A	TAGTCTTATTGCAAAATGGTGGGG	AAGGCTACTTTCTTCCCATGTC	valid
scaffold12	4998223	G	A	GAGTGTAAACCCTCCCTTAAAGGG	CTTTCCCTGTTTGGACATCACC	valid
scaffold16	3726405	A	G	CAGACCTTCTGCTTGTCTATCT	GTTCTGTCACTGGTACAGCCTGA	valid
scaffold146	3758077	A	G	TGAGGCTTATGCGCTAGATGATT	CCATGGAAGTACTCCCTCTCTA	valid
scaffold102	8068921	G	A	CCTCTTCCCTTAATGAAAGCCTA	TGCTGTGATGATTCTTAGGT	valid
scaffold102	2843559	T	A	GTTTTCTGGTGGTTTTTCTAGTG	TCAAGATTATGCCAATAGTCCA	valid
scaffold95	7648771	A	C	GGAGAGAACCATCAGGAAACTT	CTGCCGTAGCCCCACTAGAGTA	valid
scaffold104	2630873	T	C	CAACAATAATGTCTACATCCGAA	CCAGCGATGACAATGGATTATTT	valid
scaffold16	5243011	T	C	TGAGACACTTTTGGCAATCTGAA	CGAGGTCCTGAGAGGTGAAATAA	valid
scaffold16	11947256	G	A	ATATGAGGGTTTGGAAAGGGGAG	TAGGCAGAATCTAAACCCCAACA	valid
scaffold102	304844	A	G	GAGGCACCAGGTGATGTTATTTCC	CTGATTGTCAATCCTGTGATTC	valid
scaffold16	10137923	A	G	AATCCTCTAATCCAGTGTGCCAA	TCAGCCAGCATGTTTTATTTT	valid
scaffold5	25758918	C	T	GGAAATCCATTACACATCCCTGAC	CAAGTGTACCAGACCAAGAAGGG	valid
scaffold5	8573437	C	T	ATTCAGTGAGAACGCTGACCTTC	TCATCTCATCTCAGTTGCCCAA	valid
scaffold122	5172507	C	T	CACAATAAATGAAGGGGCAAAAAG	TATCAGTATTTAGGTGGGGGAA	valid
scaffold139	3125469	A	G	GCCCTTCACTGTCTCTCTGTA	GAGAGTCTCCAGGCAATGGAATA	valid
scaffold5	12698477	A	G	CAGCCTGCAAGTTGTAAGGACTT	ATAGATGCTTTAGGATGAGGGG	valid
scaffold5	18110597	G	T	TTTATGTGCAGGACAAAAGGAGC	GGAAAGGAGCTGGTAAAACAT	valid
scaffold5	6083156	A	T	TGACATCCAATATCAACTGCCAC	TCCTTTTTCATGTTAGGGCAA	valid
scaffold33	84547	G	A	CTAAAACATCCAAGTGGACCCTG	GATCTCCAGTCTCCCTAAACC	valid
scaffold6	35325892	A	G	GAGGCAGAGAGATCGGATATGAA	GTCTTGAAAGGAGAGAGCAGTGG	valid
scaffold6	46505920	G	C	ACGCCAAAACACTGAGTGGACTG	TGTGGTTGTGAAAATCAAAGG	valid
scaffold6	33235096	T	G	CTTTCACCCAGCTCATTCTTTGT	GAGCAAGCTCTTAACCACTGGA	valid
scaffold107	4674189	G	A	TAACCTGCAGATCCCTCAAATCA	CCTCTACGCTTTCAGCGACTTC	valid
scaffold6	35426317	G	A	CAGTCAAGAACAATGAATCAC	GATGACAGTACTTGGCATGTTGC	valid
scaffold6	31042499	A	G	TGACTAGTTCAATGCTCCGTGTTG	ACCATGAAAATTCAGAAGGCTCAA	valid
scaffold33	5669683	G	T	ACCAAACCTCAAGGCCAAGATT	CACAATTCATATCTTCCCAA	valid
scaffold15	5819607	A	G	TTTGGCAGACAGAAAGACAAAA	TGCATGTGATGCACCTAACACAG	valid
scaffold15	16735786	T	C	GCATGATCTTATCAGAATCAAATGG	GCATGAATGTCAGCAGCAAAA	valid
scaffold82	8671011	T	C	TGGCTCTGAGGAGTTAGAAACA	AAAGAAACCGATGAAGAAGGAG	valid
scaffold66	5581934	T	A	GGGTCAATCTCTATTGTGCCTTG	ACGTCCTTGCAGATACCAATGT	valid

scaffold67	8172888	G	A	CCTTTGGTCCAAGTAAGAGGCTT	ATGCGTCTACAGGACAGTGACAA	valid
scaffold22	7237076	C	T	CACGATAAAATCTCCAGTGTGG	AAATTCAAAAGCAATTTGGTATGC	valid
scaffold110	2802003	A	G	CTCCAGGTGTGTTATGAGGCTCAG	GCAGGCTGGTCAGCTCACTTAC	valid
scaffold22	3960128	G	A	CAGTTAGGGATGGGTGAAACAG	ATCCATGCTGCATTCTAAGCATT	valid
scaffold67	10022414	A	T	CATTGCTAGGATACCTGGCTGAG	ACCCGGGTACTTAGTGATCATT	valid
scaffold82	7471496	T	G	TCACATGGATACCTTCTCCAAGA	TGGTCACTTACTTGGCTAGCTT	valid
scaffold22	16225539	C	T	GCTTTGTGCAGTCACAAGCTCTA	GCTTCAGTAGAAGGCAGGCATTA	valid
scaffold67	9313089	G	A	AAACTCACCTTTCAACAAATGCC	AAAGGGGTAAGGAAATTTGACCA	valid
scaffold25	10103590	A	G	TTCTCCATGCTCACATACCTTT	TGTGCCATTTTCTTACGCTCAA	valid
scaffold40	12410382	G	C	GCCTCCTTTACTTGTCTCCATT	AGCCCTGGCTTAAACAGAAATA	valid
scaffold130	2242783	C	T	ATCCAGTGGTGCAAAATATCAAGG	GGACCCCTGTAAGATAAATGCT	valid
scaffold25	11661341	G	A	TGCAGATATGCCTAATAACCC	TTCAGATCATTCTGCAGTTCAA	valid
scaffold34	4386574	A	G	TTATGCCAAAAACCTGGAACATT	ACAGATGCTCACCAAGATGCTG	valid
scaffold134	1751683	A	C	CAGCCACTCTGCCTTATGTTTCT	TGTGTGAAGCTGCCATTAAGGA	valid
scaffold56	16255914	C	A	TGCAAAATGGTAGGATCTGACTGA	CTGTGGCAATCTGACAAAAGTTG	valid
scaffold56	53938	G	A	CTGACACAAAAGTTGGGAAAAGG	TCTCTGTCCAAAACAAATGCTCAA	valid
scaffold219	1621319	T	C	ACTTCATCCTCACGATAGCCTTG	AGGACTCCCAGTGTCTTATC	valid
scaffold56	12034959	C	T	GACCAGCTCAGAAATGAGGATGT	CACACTGTTTCCCATAGAGGGAC	valid
scaffold56	13635220	T	A	AGCAGGCAATCAAAGTGCTTAAAC	ACAGGGAAGTTGACAGGATCAAAA	valid
scaffold123	6899529	G	C	CCTATTGTGCACCCATTACCATTG	TTTGCCTCTCTCTCGTGGAAATA	valid
scaffold92	6593730	A	G	ACTTTAGGGATCTGGGTCTCTGG	ATGAAGCCCTATTTTTCTGTCTC	valid
scaffold219	672728	C	T	GATCATCTGCATGTGGTGAGAAC	TTTCAAACCTGGGAAAACATTGCT	valid
scaffold153	404406	G	A	CAGGGCCTTTCCAGGAGATAGAAAT	TTAGCTGAGACATCAGTGCCTTG	valid
scaffold20	284659	A	C	CCTGAACATGGTGCATTACAAAA	TTGGTATTTGCAAGGCTATTGCT	valid
scaffold87	5452327	T	C	GTCTGTCTCTGGCAATTGAGIT	CCAAGTGATTGCTACTCAACAG	valid
scaffold32	11004997	T	G	AACATTTGCAATGAAGCCATCTT	TTCTTAGTCCCCATCATACCT	valid
scaffold32	799857	C	A	ACTCTGTTGCCCTAAGACTTGTG	GAATTGTGATGCCATCTAGGACC	valid
scaffold20	10935955	G	T	CAAGACCATTGGAAGGGATAACA	TACAGCTGCTTTTTCTGTAGGG	valid
scaffold178	2695746	G	A	ATGGTTACAAAACCAATCACTC	TGAATTCAGATTATGCCCTTGG	valid
scaffold3	11133200	G	A	TGATTGCCCTTAAAAATGGTTACGC	GCTGGCATTCCAAAAATAGTGC	valid
scaffold38	8794148	A	G	AGGGTTGTGCTTCTAATGGTGA	TGCTCTGCATAACAGTTTGTGCT	valid
scaffold8	16719278	C	T	GGGAACCATGCTTCTTCTATCT	CAAAAGACAGCCGATGTCAGTG	valid
scaffold9	7008210	C	A	CTCAGAAGATGAGATGGGCTGAT	TCAGGTTCTTGGAACTCCATGT	valid
scaffold3	18069101	A	C	AGCACTCCCTCACCTATTCTTTC	TGTGGCTCAGATAACAGGTCAA	valid
scaffold38	1352310	A	C	TAGAAGAAGTCAGGATGCCAAGC	AAAACCTGGAATTTGGGAGGTA	valid
scaffold99	3450937	G	A	ACCCAGCATATTCAGGGGATATT	GGATTGCATTTGCAGTTTGTCT	valid
scaffold48	7878594	T	C	CATTTTGTGGCACTGTCTTTGAG	AGGAGTCTCTCTTTTCTCCCA	valid
scaffold99	5562915	T	G	CACACTGAGGTGAGTCTTCTT	TGCTGTCTCCATGACACATCCT	valid
scaffold41	6377180	G	T	TAGTAGTATGCGAAAGGGGCTCA	ATTGGCATGTGCTCACTCAATAA	valid
scaffold41	6377181	A	T	TAGTAGTATGCGAAAGGGGCTCA	ATTGGCATGTGCTCACTCAATAA	valid
scaffold59	9934332	C	T	CCCTTACGTGATCAGACCAATAGG	TTGCCCTCAATAGATACCAGCAT	valid
scaffold23	878261	T	G	CCTCATGAGAACCTCAGAAAGGAA	GGCTTTGTCTAAGCTGTGCTCA	valid
scaffold101	620502	G	A	AATCCCTCCTCCATGAAGTTTTC	ACCTACTAAAGGCCACGTGTCA	valid
scaffold180	2378405	T	C	TTGCCAAGTGTACCCTAAGGAT	TGCCTAACATCACCAAGTGAAGAA	valid
scaffold116	4348238	G	T	GCTCAGAACCCGTCTCATCTTTA	GCTGACTACTCTTCCATTGCC	valid
scaffold17	10419245	T	C	CAACCTCTCTTGTCTCTGGA	CCCTTGCCTTTCTGTATAACT	valid
scaffold57	11179091	A	G	GTGGTGTAGCCAAAGTTGGTAAGG	CAGGGCTTTTGTGTTTTCTCAAC	valid
scaffold57	11179169	G	T	GTGGTGTAGCCAAAGTTGGTAAGG	CAGGGCTTTTGTGTTTTCTCAAC	valid
scaffold1	1765629	G	A	AACTCTGAGCCTGTCTTCCAG	TGCATGCTCTCTGGGTTTATTT	valid
scaffold5	8573549	T	C	ATTCACTGAGAACGCTGACCTTC	TCATCTCATCTCAGTTGCCCAA	valid
scaffold5	12698633	T	C	CAGCCTGAGTTGTAAGGACTTT	ATAGATGCTTTAGGATGAGGGGC	valid
scaffold6	46505955	T	C	ACGCCAAAACCTACTGGATGACTG	TGTGGTTGTGAAAAATCAAAAAGG	valid
scaffold6	33235136	C	T	CTTTCACCCAGCTCATTCTTTGT	GAGCAAGCTTAAACCACTGGA	valid
scaffold66	5581761	T	A	GGGTCATTCTCTATTGTGCCTTG	ACGTCCTTGCAGATACCAAATGT	valid
scaffold66	5581789	A	C	GGGTCATTCTCTATTGTGCCTTG	ACGTCCTTGCAGATACCAAATGT	valid
scaffold66	5581896	A	G	GGGTCATTCTCTATTGTGCCTTG	ACGTCCTTGCAGATACCAAATGT	valid
scaffold66	5581964	A	C	GGGTCATTCTCTATTGTGCCTTG	ACGTCCTTGCAGATACCAAATGT	valid
scaffold66	5581973	A	C	GGGTCATTCTCTATTGTGCCTTG	ACGTCCTTGCAGATACCAAATGT	valid
scaffold22	3960249	T	C	CAGTTAGGGATGGGTGAAACAG	ATCCATGCTGCATTCTAAGCATT	valid
scaffold22	3960291	T	C	CAGTTAGGGATGGGTGAAACAG	ATCCATGCTGCATTCTAAGCATT	valid
scaffold130	2242802	A	G	ATCCAGTGGTGCAAAATATCAAGG	GGACCCCTGTAAGATAAATGCT	valid
scaffold56	16255997	T	C	TGCAAAATGGTAGGATCTGACTGA	CTGTGGCAATCTGACAAAAGTTG	valid
scaffold219	672798	A	G	GATCATCTGCATGTGGTGAGAAC	TTTCAAACCTGGGAAAACATTGCT	valid
scaffold20	284698	T	G	CCTGAACATGGTGCATTACAAAA	TTGGTATTTGCAAGGCTATTGCT	valid
scaffold178	2695816	T	C	ATGGTTACAAAACCAATCACTC	TGAATTCAGATTATGCCCTTGG	valid
scaffold178	2695863	C	A	ATGGTTACAAAACCAATCACTC	TGAATTCAGATTATGCCCTTGG	valid
scaffold17	10419245	T	C	CAACCTCTCTTGTCTCTGGA	CCCTTGCCTTTCTGTATAACT	valid
scaffold16	3726386	A	G	CAGACCTTCTGCCTTGTCTCTT	GTCTGTCACTGGTACAGCCTGA	valid
scaffold99	12217751	G	T	CTAAAAATACATGGCAGGCAGC	ATTTGACTCTGGCTTCTGCATC	PCR fail
scaffold16	6801448	G	A	CCAGAGTGCAAGAATGGAGACTT	TTTCTCAAACCTGGACTGACACA	PCR fail

**Supplementary Table 16 | Scaffolds identified as putative sex chromosome regions.** The whole genome scale alignment between cattle and minke whale was performed by LASTZ. Some of minke whale scaffolds were aligned to cattle's sex chromosomes. A total eight scaffolds, which were aligned to cattle genome with >10Kb of length and >0.5 of coverage, were identified as putative sex chromosome regions of the minke whale.

Scaffold	Alignment length	Alignment coverage	Gene in scaffold
scaffold725	17,543	0.59	<i>RPS6KA6</i>
scaffold398	148,356	0.64	<i>MAGEB18, MAGEB4, MAGEB5, MAGEB5</i>
scaffold863	11,939	0.62	-
scaffold647	22,914	0.56	<i>TRO</i>
scaffold573	34,299	0.56	<i>FAM123B, ASB12</i>
scaffold352	231,012	0.56	<i>ODZI, SH2DIA</i>
scaffold745	21,083	0.72	-
scaffold634	23,433	0.53	<i>PCYT1B</i>

**Supplementary Table 17 | Statistics of the predicted protein-coding genes.** The average transcript length does not contain UTRs except that of the final gene set. Three approaches were used for gene prediction: Homolog (*Bos taurus*, *Canis familiaris*, *Homo sapiens*, *Monodelphis domestica*, *Ornithorhynchus anatinus*, *Sus scrofa*, and *Tursiops truncatus*), *De novo* (*GENSCAN*, *AUGUSTUS*), and RNA-seq (*Cufflinks*). The results on homologs and *de novo* predictions were consolidated using *GLEAN*. *Cufflinks* predictions were used to construct the final set.

	Gene set	Number	Average transcript length (bp)	Average CDS length (bp)	Average exon per gene	Average exon length (bp)	Average intron length (bp)
<i>De novo</i>	<i>AUGUSTUS</i>	20,903	48686.57	1434.22	8.8	162.91	6055.11
	<i>GENSCAN</i>	42,586	39227.16	1293.36	7.98	162.08	5434.88
Homolog	<i>B. taurus</i>	22,541	23549.89	1397.99	8.02	174.27	3154.62
	<i>C. familiaris</i>	22,402	23049.98	1368.95	7.92	172.86	3133.32
	<i>H. sapiens</i>	22,543	24922.28	1417.87	8.1	175.14	3312.47
	<i>M. domestica</i>	21,460	22571.8	1315.18	7.51	175.02	3262.9
	<i>O. anatinus</i>	21,143	17784.73	1096.91	6.27	174.97	3167.13
	<i>S. scrofa</i>	21,251	19364.86	1220.47	6.94	175.81	3053.62
	<i>T. truncatus</i>	21,459	25339.17	1409.43	8.14	173.11	3350.55
	<i>GLEAN</i>	17,531	23561.13	1469.29	8.23	178.63	3057.63
	Final	20,605	30442.56	1452.75	8.22	176.69	3638.74

**Supplementary Table 18 | Summary of evidence for the final gene models.** P: *ab initio* prediction; H: homology-based; C: supported by RNA-seq data. Based on a number of supported gene sources, the evidence was further separated into ‘single’ (with one gene source) and ‘multiple’ (with two or more gene sources). The overlap threshold was relative to the CDS region in the final gene set.

	≥20% overlap		≥50% overlap		≥80% overlap	
	Number	Percent(%)	Number	Percent(%)	Number	Percent(%)
P(single)	19	0.09	311	1.51	1,491	7.24
P(multiple)	986	4.79	1,098	5.33	1,299	6.30
H(single)	1	0.00	6	0.03	45	0.22
H(multiple)	5	0.02	29	0.14	183	0.89
C(single)	0	0.00	1	0.00	28	0.14
C(multiple)	0	0.00	0	0.00	0	0.00
H+C	4	0.02	33	0.16	307	1.49
P+C	164	0.80	236	1.15	598	2.90
P+H	4,147	20.13	4,352	21.12	4,043	19.62
P+H+C	12,199	59.20	11,406	55.36	9,149	44.40

**Supplementary Table 19 | Functional annotation statistics.** Predicted gene functions were assigned based on the best matches in the alignments using BLASTP with the InterPro, GO, KEGG, SwissProt, and TrEMBL databases.

		Number	Percent (%)
Total		20,605	
Annotated	InterPro	16,005	77.68
	GO	13,202	64.07
	KEGG	14,256	69.19
	SwissProt	18,586	90.20
	TrEMBL	18,776	91.12
Not annotated		1,805	8.76

**Supplementary Table 20 | miRNA, snRNA, tRNA, and rRNA statistics.** Four types of non-coding RNAs (ncRNAs) in the minke whale genome were annotated. miRNA is micro RNA; tRNA is transfer RNA; rRNA is ribosomal RNA; snRNA is small nuclear RNA.

Type	Copy Number	Average length (bp)	Total length (bp)	% of genome	
miRNA	660	86.5	57,088	0.0023	
tRNA	317	76.63	24,292	0.001	
rRNA	rRNA total	325	109.39	35,553	0.0015
	18S	47	200.04	9,402	0.0004
	28S	123	126.80	15,597	0.0006
	5.8S	3	100.67	302	0.0000
	5S	152	67.45	10,252	0.0004
snRNA	snRNA total	1,296	113.37	146,930	0.006
	CD-box	304	88.01	26,755	0.0011
	HACA-box	288	135.34	38,979	0.0016
	Unknown	18	31.28	563	0.0000
	scaRNA	19	156.63	2,976	0.0001
	splicing	667	116.43	77,657	0.0032

**Supplementary Table 21 | Summary of repeats in the minke whale genome.** Repetitive elements were searched in the minke whale genome. Repetitive elements occupied 37.3% of the minke whale genome.

Type	Repeat Size (bp)	% of genome
TRF	35,678,411	1.46
RepeatMasker	860,676,796	35.23
RepeatProteinMask	249,732,325	10.22
<i>De novo</i>	720,787,721	29.51
Total	912,163,603	37.34

**Supplementary Table 22 | Repeat statistics for the minke whale genome.** Repbase TEs: results of RepeatMasker analysis using Repbase; TE proteins: results of RepeatProteinMask analysis using Repbase; *De novo*: results of RepeatMasker analysis using the library predicted by the *de novo* method; Combined: the combined results for Repbase TEs, TE proteins, and the *de novo* method.

	RepBase TEs		TE Proteins		<i>De novo</i>		Combined TEs	
	Length (bp)	% in Genome	Length (bp)	% in Genome	Length (bp)	% in Genome	Length (bp)	% in Genome
DNA	77,730,528	3.18	8,675,160	0.36	48,229,557	1.97	82,348,333	3.37
LINE	508,513,598	20.82	230,902,172	9.45	461,168,564	18.88	557,302,190	22.83
LTR	135,732,724	5.56	10,185,678	0.42	97,499,816	3.99	140,879,797	5.77
SINE	143,296,510	5.87	0	0	115,299,510	4.72	149,734,808	6.13
Other	1,955	0.00	0	0	0	0	1,955	0.00
Unknown	0	0	0	0	10,079,967	0.41	10,079,967	0.41
Total	860,676,796	35.23	249,732,325	10.22	720,336,885	29.49	897,120,885	36.72

**Supplementary Table 23 | Mammalian repeat element statistics for the minke whale genome**

Type	Number of copies	Length occupied (bp)	Fraction of genome (%)
<b>LINEs</b>	1,014,625	479,878,285	19.6438
L1	898,051	451,370,901	18.4769
RTEX	8,332	2,270,414	0.0929
RTEX	138	25,920	0.0011
CR1	107,870	26,146,438	1.0703
R4	234	64,612	0.0026
<b>SINEs</b>	706,757	142,611,670	5.8378
tRNA	706,358	142,552,806	5.8354
7SL	38	7,204	0.0003
5S	361	51,660	0.0021
<b>LRT retrotransposons</b>	236,416	76,057,931	3.1134
Gypsy	13,088	1,532,854	0.0627
ERV1	99,887	32,796,200	1.3425
ERV2	13,006	997,310	0.0408
ERV3	110,435	40,731,567	1.6673
<b>DNA transposons</b>	320,039	67,084,863	2.7461
hAT	240,961	46,093,194	1.8868
Mariner/Tc1	68,667	19,798,220	0.8104
MuDR	4,078	398,147	0.0163
PiggyBac	482	89,567	0.0037
Helitron	5,851	705,735	0.0289
<b>Others</b>	27	2,040	0.0000
<b>Total</b>	<b>2,277,864</b>	<b>765,634,789</b>	<b>31.3413</b>

**Supplementary Table 24 | Composition of the transposable elements in the minke whale genome identified based on *de novo* RNA sequencing.** Transposable elements in the minke whale genome were identified using the assembled transcripts. Out of three muscle tissues, one was used in this analysis.

	Fraction of <i>de novo</i> RNA sequencing (%)					
	Brain	Heart	Kidney	Liver	Lung	Muscle
<b>LINEs</b>	4.5647	3.2818	3.9101	4.2107	3.3529	2.8381
L1	2.8252	2.0225	2.4357	2.7640	2.0705	1.8427
RTEX	0.0603	0.0277	0.0380	0.0371	0.0319	0.0313
RTE	0.0003	0.0006	0.0007	0.0007	0.0006	0.0007
CR1	1.6774	1.2293	1.4338	1.4074	1.2486	0.9621
R4	0.0015	0.0018	0.0018	0.0014	0.0012	0.0012
<b>SINEs</b>	1.5310	1.3107	1.4531	1.5388	1.2701	1.2152
tRNA	1.5275	1.3059	1.4495	1.5328	1.2671	1.2110
5S	0.0035	0.0048	0.0036	0.0060	0.0030	0.0042
<b>LTR retrotransposons</b>	1.1676	1.0158	1.2684	1.2433	0.9774	0.8854
Gypsy	0.0242	0.0198	0.0380	0.0230	0.0184	0.0133
ERV1	0.2301	0.2026	0.2532	0.2560	0.1776	0.2084
ERV2	0.0006	0.0013	0.0009	0.0007	0.0006	0.0035
ERV3	0.9126	0.7921	0.9763	0.9635	0.7808	0.6601
<b>DNA transposons</b>	1.2170	0.9908	1.1758	1.4093	1.0934	0.9231
hAT	0.9019	0.7203	0.8451	0.9919	0.7530	0.6580
Mariner/Tc1	0.3126	0.2682	0.3281	0.4141	0.3377	0.2634
PiggyBac	0.0025	0.0023	0.0025	0.0033	0.0026	0.0017
<b>Others</b>	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
<b>Total</b>	8.4903	6.6090	7.8174	8.4121	6.7038	5.8718

**Supplementary Table 25 | Statistics for the orthologous gene families.** Unclustered genes refer to those specific to the current species; Unique families refer to those specific to the current species.

Species	#Total genes	#Genes in families	#Unclustered genes	#Families	#Unique families	Ave. genes per family
<i>B. acutorostrata</i>	20,605	17,950	2,655	12,675	46	1.42
<i>T. truncatus</i>	16,515	16,276	239	12,678	2	1.28
<i>B. Taurus</i>	19,970	19,795	175	13,288	8	1.49
<i>S. scrofa</i>	21,577	20,029	1,548	13,192	79	1.52
<i>C. familiaris</i>	19,839	19,022	817	13,256	17	1.43
<i>H. sapiens</i>	22,292	21,463	829	13,520	95	1.59
<i>M. domestica</i>	21,299	20,273	1,026	12,813	65	1.58
<i>O. anatinus</i>	21,663	16,121	5,542	13,669	530	1.18

**Supplementary Table 26 | Summary of segmental duplications in the minke whale genome.** The segmental duplications in the minke whale genome were detected with different cutoff sizes (1Kb, 5Kb, 10Kb, and 50Kb). When we used the >1Kb cutoff, a total 33.4Mb of segmental duplication regions were detected.

Cutoff	# block	Median size (bp)	Genome coverage (bp)
>1kb	11,701	1,749	33,429,537
>5kb	741	6,223	7,219,683
>10kb	76	11,082	1,420,861
>50kb	0	0	0

**Supplementary Table 27 | Parameters used in LASTZ.** The parameters not shown in this table were the default values used by LASTZ.

Target	Query	T	C	H	Y	L	K
<i>B. acutorostrata</i>	<i>S. scrofa</i>	2	2	2000	3400	6000	2200
<i>B. acutorostrata</i>	<i>T. truncatus</i>	2	2	2000	3400	6000	2200
<i>S. scrofa</i>	<i>T. truncatus</i>	2	2	2000	3400	6000	2200

**Supplementary Table 28 | Statistics of the aligned genomes.** The genomes were hard masked with Ns in repeat sequence regions prior to LASTZ alignment.

Genome	#scaffolds	Genome size (bp, with N)	Genome Size (bp, without N)	Genome Size (bp, masked)	%masked
<i>B. acutorostrata</i>	104,325	2,442,893,294	2,297,918,384	1,387,065,169	37.29
<i>S. scrofa</i>	20	2,262,596,414	2,231,331,862	1,299,683,467	41.18
<i>T. truncatus</i>	111,212	2,521,907,548	2,298,444,090	1,300,529,065	39.57

**Supplementary Table 29 | Genome alignment statistics.** The minke whale's genomic synteny with bottlenose dolphin and pig were detected by LASTZ whole genome alignment program.

Species vs Species	Aligned Length (bp)	Target Genome Coverage Rate (%)	Query Genome Coverage Rate (%)
<i>B. acutorostrata</i> vs <i>S. scrofa</i>	681,551,392	28.07	30.30
<i>B. acutorostrata</i> vs <i>T. truncatus</i>	1129,860,272	47.17	45.70
<i>S. scrofa</i> vs <i>T. truncatus</i>	840,246,530	37.14	33.32

**Supplementary Table 30 | Annotated domains of cetacean-specific (the minke whale and bottlenose dolphin) protein families.** P-value is calculated by Fisher's exact test with a conservative 5% false discovery rate criterion.

InterPro ID	Description	# of domains shared only between minke whale and bottlenose dolphin	# of domains shared only between cow and pig	P-value
<b>Immunity and host defense</b>				
IPR013098	Immunoglobulin I-set	73	25	2.20E-16
IPR007110	Immunoglobulin-like	71	38	2.20E-16
IPR003599	Immunoglobulin subtype	78	32	2.20E-16
IPR003598	Immunoglobulin subtype 2	71	4	2.20E-16
IPR013783	Immunoglobulin-like fold	84	0	2.20E-16
<b>Sensory perception</b>				
IPR000276	GPCR, rhodopsin-like, 7TM	0	1019	2.20E-16
IPR017452	GPCR, rhodopsin-like superfamily	0	518	2.20E-16
IPR000725	Olfactory receptor	0	469	2.20E-16
IPR019424	7TM GPCR, olfactory receptor/chemoreceptor Srsx	0	218	2.20E-16
IPR007960	Mammalian taste receptor	0	20	0.003696
<b>Others</b>				
IPR001452	Src homology-3 domain	21	0	8.60E-13
IPR001609	Myosin head, motor domain	14	0	8.83E-09
IPR000884	Thrombospondin, type 1 repeat	24	9	4.42E-08
IPR003961	Fibronectin, type III	15	2	1.80E-07
IPR000998	MAM domain	9	0	6.57E-06
IPR000857	MyTH4 domain	9	0	6.57E-06
IPR002126	Cadherin	14	4	8.62E-06
IPR001849	Pleckstrin homology domain	17	8	1.80E-05
IPR002557	Chitin binding domain	8	0	2.47E-05
IPR000008	C2 calcium-dependent membrane targeting	12	4	7.33E-05
IPR001265	Formin	11	3	7.34E-05
IPR022812	Dynamin	7	0	9.29E-05
IPR000850	Adenylate kinase	7	0	9.29E-05

**Supplementary Table 31 | Annotated domains of minke whale-specific protein families.**  
P-value is calculated by Fisher's exact test with a conservative 5% false discovery rate criterion.

InerPro ID	Description	# of domains in minke whale-specific families	# of domains in bottlenose dolphin-specific families	# of domains in cow-specific families	# of domains in pig-specific families	P-value
IPR002126	Cadherin	199	4	4	7	2.20E-16
IPR001909	Krueppel-associated box	4	0	20	4	3.21E-15
IPR013783	Immunoglobulin-like fold	48	0	0	0	3.48E-15
IPR002223	Proteinase inhibitor I2, Kunitz metazoa	0	5	10	0	1.58E-13
IPR007087	Zinc finger, C2H2	59	3	8	4	5.71E-12
IPR006080	Beta defensin/Neutrophil defensin	0	0	11	2	8.85E-11
IPR008197	Whey acidic protein, 4-disulphide core	0	0	9	0	2.31E-10
IPR001855	Beta defensin type	0	0	11	4	7.10E-10
IPR013087	Zinc finger C2H2-type/integrase DNA-binding domain	32	0	0	0	7.32E-10
IPR001356	Homeodomain	20	8	0	0	3.03E-09
IPR001878	Zinc finger, CCHC-type	3	0	0	37	3.62E-09
IPR015919	Cadherin-like	43	1	1	4	1.55E-08
IPR001849	Pleckstrin homology domain	0	0	5	26	1.85E-08
IPR020894	Cadherin conserved site	28	1	0	0	2.55E-08
IPR006652	Kelch repeat type 1	33	0	0	2	4.89E-08
IPR000795	Protein synthesis factor, GTP-binding	24	0	0	0	2.81E-07
IPR002494	High sulphur keratin-associated protein	0	0	6	0	5.83E-07
IPR000558	Histone H2B	0	0	6	0	5.83E-07
IPR000719	Protein kinase, catalytic domain	17	2	2	57	1.68E-06
IPR003597	Immunoglobulin C1-set	0	2	6	2	1.96E-06
IPR001245	Serine-threonine/tyrosine-protein kinase catalytic domain	5	1	1	34	4.01E-06
IPR000294	Gamma-carboxylglutamic acid-rich (GLA) domain	0	0	5	0	4.08E-06
IPR004000	Actin-like	24	0	0	1	4.18E-06
IPR000242	Protein-tyrosine phosphatase, receptor/non-receptor type	0	0	0	19	4.34E-06
IPR009057	Homeodomain-like	27	4	1	2	4.47E-06
IPR004244	Transposase, L1	20	0	0	0	5.71E-06
IPR002348	Interleukin 1/heparin-binding growth factor	0	4	3	2	8.67E-06
IPR002172	Low-density lipoprotein (LDL) receptor class A repeat	0	2	4	0	9.61E-06
IPR000569	HECT	0	0	0	18	9.73E-06
IPR000079	High mobility group nucleosome-binding domain-containing family	0	4	0	0	1.02E-05
IPR001564	Nucleoside diphosphate kinase	0	4	0	0	1.02E-05
IPR000157	Toll/interleukin-1 receptor homology (TIR) domain	0	4	0	0	1.02E-05
IPR001452	Src homology-3 domain	15	0	0	42	1.04E-05
IPR015880	Zinc finger, C2H2-like	30	1	4	3	1.43E-05
IPR001199	Cytochrome b5	0	0	6	3	1.96E-05
IPR004018	RPEL repeat	18	0	0	0	2.44E-05

**Supplementary Table 32 | Differences in the frequencies of functional olfactory receptor genes in different species.** The number of olfactory receptor genes (functional and pseudogenes) was much lower in whales than in other mammals.

Species	Number of functional genes	Number of pseudogenes	Percentage of functional gene
Pig	1,113	188	86
Rat	1,201	292	80
Dog	872	222	80
Mouse	1,037	354	75
Zebrafish	102	35	74
Human	388	414	48
Frog	410	478	46
<b>Minke whale</b>	<b>60</b>	<b>74</b>	<b>45</b>
Pufferfish	44	54	45
Chicken	82	476	15

**Supplementary Table 33 | Numbers of olfactory receptor (OR) gene-like sequences in the minke whale genome.** A total of 60 functional olfactory receptor genes were detected in the minke whale genome.

Types of OR gene like sequences	Number of sequences (%)	Sequence type definition
Functional	60 (24)	At least 300 amino acid long with complete ORF and four conserved OR motifs
Pseudo	74 (29)	At least 300 amino acid long with incomplete ORF or lacking any of conserved OR motifs
Partial	117 (47)	100 to less than 300 amino acid long with partial containing of conserved OR motifs (minimum number = 2)
Total	251 (100)	

**Supplementary Table 34 | Enrichment of the Gene Ontology (GO) categories of expanded genes in the minke whale compared with the common whale ancestor. P-values were calculated by Fisher's exact test.**

GO ID	# of genes	P-value	Q-value	GO domain	GO description
GO:0019538	493	<1.00E-12	<1.00E-12	biological_process	protein metabolic process
GO:0044267	416	2.30E-12	1.76E-11	biological_process	cellular protein metabolic process
GO:0043232	302	3.70E-12	2.78E-11	cellular_component	intracellular non-membrane-bounded organelle
GO:0043228	302	3.70E-12	2.78E-11	cellular_component	non-membrane-bounded organelle
GO:0030529	300	7.65E-12	5.57E-11	cellular_component	ribonucleoprotein complex
GO:0003735	288	1.28E-11	8.87E-11	molecular_function	structural constituent of ribosome
GO:0005840	278	1.31E-11	9.04E-11	cellular_component	Ribosome
GO:0044446	230	1.01E-10	6.65E-10	cellular_component	intracellular organelle part
GO:0016681	11	8.76E-10	5.49E-09	molecular_function	oxidoreductase activity, acting on diphenols and related substances as donors, cytochrome as acceptor
GO:0008121	11	8.76E-10	5.49E-09	molecular_function	ubiquinol-cytochrome-c reductase activity
GO:0044281	182	1.26E-09	7.87E-09	biological_process	small molecule metabolic process
GO:0006807	261	5.00E-09	3.08E-08	biological_process	nitrogen compound metabolic process
GO:0034641	252	3.39E-08	2.04E-07	biological_process	cellular nitrogen compound metabolic process
GO:0016491	162	3.49E-07	2.03E-06	molecular_function	oxidoreductase activity
GO:0006139	229	1.22E-06	6.87E-06	biological_process	nucleobase-containing compound metabolic process
GO:1901360	246	1.58E-06	8.80E-06	biological_process	organic cyclic compound metabolic process
GO:0010556	244	1.61E-06	8.86E-06	biological_process	regulation of macromolecule biosynthetic process
GO:0009069	16	1.67E-06	9.16E-06	biological_process	serine family amino acid metabolic process
GO:0009889	244	1.75E-06	9.56E-06	biological_process	regulation of biosynthetic process
GO:0046483	241	2.43E-06	1.32E-05	biological_process	heterocycle metabolic process
GO:0016675	15	5.44E-06	2.81E-05	molecular_function	oxidoreductase activity, acting on a heme group of donors
GO:0006725	238	5.54E-06	2.85E-05	biological_process	cellular aromatic compound metabolic process
GO:0010468	240	7.48E-06	3.76E-05	biological_process	regulation of gene expression
GO:0071704	250	1.44E-05	7.04E-05	biological_process	organic substance metabolic process
GO:0060255	245	1.72E-05	8.32E-05	biological_process	regulation of macromolecule metabolic process
GO:0046914	362	2.07E-05	9.98E-05	molecular_function	transition metal ion binding
GO:0051920	7	2.95E-05	1.39E-04	molecular_function	peroxiredoxin activity
GO:0005000	5	2.98E-05	1.40E-04	molecular_function	vasopressin receptor activity
GO:2001141	230	3.96E-05	1.83E-04	biological_process	regulation of RNA biosynthetic process
GO:0019222	253	3.44E-04	1.39E-03	biological_process	regulation of metabolic process
GO:0008137	8	5.59E-04	2.20E-03	molecular_function	NADH dehydrogenase (ubiquinone) activity
GO:0016655	8	5.59E-04	2.20E-03	molecular_function	oxidoreductase activity, acting on NADH or NADPH, quinone or similar compound as acceptor
GO:0043234	192	6.38E-04	2.50E-03	cellular_component	protein complex
GO:0008270	304	2.29E-03	7.86E-03	molecular_function	zinc ion binding
GO:0006096	12	4.43E-03	1.42E-02	biological_process	Glycolysis
GO:0000166	397	1.03E-02	2.80E-02	molecular_function	nucleotide binding
GO:1901265	397	1.03E-02	2.80E-02	molecular_function	nucleoside phosphate binding
GO:0016209	13	1.03E-02	2.80E-02	molecular_function	antioxidant activity

**Supplementary Table 35 | Enrichment of the Gene Ontology (GO) categories of contracted genes in the minke whale compared with the common whale ancestor. P-values were calculated by Fisher's exact test.**

GO ID	# of genes	P-value	Q-value	GO domain	GO description
GO:0018149	6	6.88E-08	1.09E-06	biological_process	peptide cross-linking
GO:0016755	6	3.64E-07	5.67E-06	molecular_function	transferase activity, transferring amino-acyl groups
GO:0000786	19	1.29E-06	1.93E-05	cellular_component	Nucleosome
GO:0006325	23	6.18E-06	9.09E-05	biological_process	chromatin organization
GO:0060089	67	6.43E-06	9.11E-05	molecular_function	molecular transducer activity
GO:0005215	74	6.97E-06	9.70E-05	molecular_function	transporter activity
GO:0046982	18	7.28E-06	9.95E-05	molecular_function	protein heterodimerization activity
GO:0006952	11	1.02E-05	1.35E-04	biological_process	defense response
GO:0006334	19	1.43E-05	1.74E-04	biological_process	nucleosome assembly
GO:0004252	18	1.59E-05	1.91E-04	molecular_function	serine-type endopeptidase activity
GO:0051276	23	2.31E-05	2.73E-04	biological_process	chromosome organization
GO:0044427	20	3.04E-05	3.52E-04	cellular_component	chromosomal part
GO:0019200	5	3.07E-05	3.52E-04	molecular_function	carbohydrate kinase activity
GO:0015297	8	3.49E-05	3.94E-04	molecular_function	antiporter activity
GO:0040007	7	1.01E-04	1.02E-03	biological_process	Growth
GO:0017171	18	1.09E-04	1.08E-03	molecular_function	serine hydrolase activity
GO:0016020	116	1.25E-04	1.19E-03	cellular_component	Membrane
GO:0004175	31	2.45E-04	2.03E-03	molecular_function	endopeptidase activity
GO:0004872	60	2.72E-04	2.14E-03	molecular_function	receptor activity
GO:0008021	6	2.72E-04	2.14E-03	cellular_component	synaptic vesicle
GO:0038023	53	2.83E-04	2.21E-03	molecular_function	signaling receptor activity
GO:0055085	43	3.13E-04	2.42E-03	biological_process	transmembrane transport
GO:0008233	40	3.17E-04	2.42E-03	molecular_function	peptidase activity
GO:0016459	8	3.74E-04	2.71E-03	cellular_component	myosin complex
GO:0070011	38	4.63E-04	3.28E-03	molecular_function	peptidase activity, acting on L-amino acid peptides
GO:0046983	27	6.23E-04	4.19E-03	molecular_function	protein dimerization activity
GO:0015298	5	6.68E-04	4.45E-03	molecular_function	solute:cation antiporter activity
GO:0015075	47	7.24E-04	4.63E-03	molecular_function	ion transmembrane transporter activity
GO:0022857	54	7.67E-04	4.74E-03	molecular_function	transmembrane transporter activity
GO:0022804	20	1.07E-03	6.27E-03	molecular_function	active transmembrane transporter activity
GO:0006471	3	1.11E-03	6.46E-03	biological_process	protein ADP-ribosylation
GO:0005886	12	1.15E-03	6.64E-03	cellular_component	plasma membrane
GO:0015079	13	1.17E-03	6.71E-03	molecular_function	potassium ion transmembrane transporter activity
GO:0019882	4	1.19E-03	6.77E-03	biological_process	antigen processing and presentation
GO:0051234	103	1.27E-03	7.12E-03	biological_process	establishment of localization
GO:0022891	49	1.35E-03	7.52E-03	molecular_function	substrate-specific transmembrane transporter activity
GO:0004984	17	1.36E-03	7.52E-03	molecular_function	olfactory receptor activity
GO:0005261	18	1.42E-03	7.74E-03	molecular_function	cation channel activity
GO:0005516	5	1.44E-03	7.74E-03	molecular_function	calmodulin binding
GO:0043231	84	1.47E-03	7.74E-03	cellular_component	intracellular membrane-bounded organelle
GO:0005267	12	1.62E-03	8.42E-03	molecular_function	potassium channel activity
GO:0005057	6	1.65E-03	8.52E-03	molecular_function	receptor signaling protein activity
GO:0046873	24	1.75E-03	8.78E-03	molecular_function	metal ion transmembrane transporter activity
GO:0005216	26	1.75E-03	8.78E-03	molecular_function	ion channel activity
GO:0022836	21	1.82E-03	8.92E-03	molecular_function	gated channel activity
GO:0015267	26	1.85E-03	8.95E-03	molecular_function	channel activity

**Supplementary Table 36 | Pathway analysis of the function altering genes in whale-specific lineage.** Function altering amino acid changes were predicted using PolyPhen2 (probably or possibly damaging).

KEGG pathway name	Genes having whale lineage specific amino acid changes, causing functional change
ABC transporters	<i>ABCB8,CFTR,ABCB6,ABCB5,ABCA5</i>
ECM-receptor interaction	<i>IBSP,TNC,TNN,ITGA3,COL5A3,SDC2</i>
Lysine degradation	<i>SETDB1,PLOD3,AASS</i>
Pantothenate and CoA biosynthesis	<i>ENPP1,VNN1</i>
Apoptosis	<i>PIK3R5,BIRC2,IL1A,ATM,PIK3R2</i>
Small cell lung cancer	<i>PIK3R5,ITGA3,BIRC2,TRAF3,PIK3R2</i>
Protein digestion and absorption	<i>MEP1A,ACE2,COL5A3</i>
Jak-STAT signaling pathway	<i>IL12RB2,SPRY1,IL22RA1,EP300,IL4R,IL9,CREBBP,CSF3R,PIK3R5,IL10,PIK3R2</i>
Neurotrophin signaling pathway	<i>YWHAG,PRDM4,PIK3R5,ABL1,PIK3R2</i>
Asthma	<i>IL9,IL10</i>
mTOR signaling pathway	<i>PIK3R5,RICTOR,PIK3R2</i>
Phosphatidylinositol signaling system	<i>PIKFYVE,PIK3R5,DGKH,PIK3R2</i>
Hematopoietic cell lineage	<i>CD8B,IL4R,CSF3R,ITGA3,IL1A</i>
NOD-like receptor signaling pathway	<i>MEFV,ERBB2IP,TNFAIP3,BIRC2</i>
Type II diabetes mellitus	<i>GCK,PIK3R5,PIK3R2</i>
Cell cycle	<i>YWHAG,EP300,DBF4,CREBBP,ESPL1,ABL1,ATM</i>
Vitamin digestion and absorption	<i>MMACHC,SLC46A1</i>
Peroxisome	<i>PEX6,PEX5</i>
Steroid biosynthesis	<i>FDFT1</i>
Renin-angiotensin system	<i>ACE2</i>

**Supplementary Table 37 | Positively selected genes (PSGs) in the minke whale.** The minke whale was used as the foreground branch and the cow and pig were used as the background branches for the PSGs of the minke whale. PSGs were selected by a conservative 10% false discovery rate criterion.

Protein ID	Gene Name	P-value	Q-value	Protein ID	Gene Name	P-value	Q-value
BACU010038	<i>NEB</i>	<1.00E-11	<1.00E-08	BACU017549	<i>RG9MTD3</i>	2.86E-03	3.55E-02
BACU017973	<i>CUL9</i>	<1.00E-11	<1.00E-08	BACU004164	<i>C4orf49</i>	2.96E-03	3.65E-02
BACU017034	<i>Colla2</i>	<1.00E-11	<1.00E-08	BACU012842	<i>QSOX1</i>	3.14E-03	3.83E-02
BACU009279	<i>RRBP1</i>	3.71E-11	1.62E-08	BACU004377	<i>RRP1B</i>	3.15E-03	3.83E-02
BACU013479	<i>PKDREJ</i>	2.76E-09	9.67E-07	BACU000324	<i>MUTYH</i>	3.22E-03	3.88E-02
BACU004085	<i>SLC25A48</i>	1.35E-07	3.94E-05	BACU012077	<i>AGFG2</i>	3.23E-03	3.88E-02
BACU015015	<i>RANBP3L</i>	2.30E-07	5.43E-05	BACU008106	<i>KIAA0748</i>	3.38E-03	4.03E-02
BACU001186	<i>Till10</i>	2.48E-07	5.43E-05	BACU011954	<i>PRSS53</i>	3.49E-03	4.13E-02
BACU016797	<i>KIDINS220</i>	3.66E-07	7.12E-05	BACU000024	<i>SLC30A1</i>	3.51E-03	4.13E-02
BACU009769	<i>MAGIX</i>	9.90E-07	1.73E-04	BACU000837	<i>Ankrd54</i>	3.60E-03	4.20E-02
BACU005801	<i>DHX9</i>	1.29E-06	2.05E-04	BACU003205	<i>AP4E1</i>	3.68E-03	4.27E-02
BACU002716	<i>CEPT1</i>	1.49E-06	2.08E-04	BACU008649	<i>HOMER2</i>	3.78E-03	4.34E-02
BACU018707	<i>Tmprss11f</i>	1.54E-06	2.08E-04	BACU017185	<i>IRS4</i>	3.79E-03	4.34E-02
BACU019790	<i>LRRC56</i>	2.49E-06	3.06E-04	BACU012027	<i>CCDC15</i>	4.11E-03	4.67E-02
BACU009248	<i>Fbxl16</i>	2.62E-06	3.06E-04	BACU020381	<i>OPALIN</i>	4.13E-03	4.67E-02
BACU017671	<i>SMAD4</i>	3.29E-06	3.60E-04	BACU017109	<i>Dgcr14</i>	4.19E-03	4.71E-02
BACU009355	<i>FSCB</i>	3.66E-06	3.77E-04	BACU014408	<i>CFI</i>	4.25E-03	4.72E-02
BACU014530	<i>GPR113</i>	5.06E-06	4.92E-04	BACU013112	<i>Tcof1</i>	4.26E-03	4.72E-02
BACU000176	<i>MPHOSPH8</i>	5.85E-06	5.39E-04	BACU015178	<i>HRH1</i>	4.28E-03	4.72E-02
BACU002214	<i>MUC7</i>	6.42E-06	5.62E-04	BACU009672	<i>SH3D21</i>	4.32E-03	4.72E-02
BACU020345	<i>C4BPA</i>	7.13E-06	5.95E-04	BACU010884	<i>DKC1</i>	4.38E-03	4.72E-02
BACU003650	<i>NMBR</i>	7.72E-06	6.15E-04	BACU014313	<i>SLC18B1</i>	4.40E-03	4.72E-02
BACU010209	<i>ROBO4</i>	8.32E-06	6.34E-04	BACU010313	<i>CALR3</i>	4.43E-03	4.72E-02
BACU007541	<i>CECR2</i>	1.07E-05	7.81E-04	BACU011042	<i>SEPT2</i>	4.45E-03	4.72E-02
BACU007225	<i>RPRD2</i>	1.32E-05	9.25E-04	BACU012265	<i>IFT80</i>	4.45E-03	4.72E-02
BACU016007	<i>SPINK5</i>	1.55E-05	1.04E-03	BACU011722	<i>RNASET2</i>	4.53E-03	4.75E-02
BACU009433	<i>CALB1</i>	1.69E-05	1.10E-03	BACU004186	<i>MFSD8</i>	4.55E-03	4.75E-02
BACU019534	<i>DPM1</i>	2.02E-05	1.23E-03	BACU005290	<i>FIGF</i>	4.56E-03	4.75E-02
BACU008820	<i>KIAA1671</i>	2.03E-05	1.23E-03	BACU015010	<i>WDR70</i>	4.58E-03	4.75E-02
BACU014374	<i>CENPE</i>	2.52E-05	1.42E-03	BACU018332	<i>RAB9A</i>	4.71E-03	4.85E-02
BACU016665	<i>PDE4C</i>	2.52E-05	1.42E-03	BACU003321	<i>SMTN</i>	4.78E-03	4.90E-02
BACU006816	<i>CCDC136</i>	2.97E-05	1.63E-03	BACU006241	<i>ERCC8</i>	4.83E-03	4.92E-02
BACU017756	<i>IL1A</i>	4.12E-05	2.19E-03	BACU004096	<i>MYOT</i>	4.89E-03	4.95E-02
BACU018907	<i>C2orf72</i>	4.64E-05	2.29E-03	BACU003633	<i>CDHR5</i>	4.93E-03	4.96E-02
BACU004347	<i>COL6A1</i>	4.96E-05	2.29E-03	BACU000128	<i>GTF3A</i>	4.95E-03	4.96E-02
BACU003138	<i>C7orf13</i>	5.03E-05	2.29E-03	BACU010633	<i>DMP1</i>	5.15E-03	5.13E-02
BACU011401	<i>FOXH1</i>	5.13E-05	2.29E-03	BACU013395	<i>SGSM3</i>	5.33E-03	5.28E-02
BACU009684	<i>SNIP1</i>	5.14E-05	2.29E-03	BACU004311	<i>C16orf95</i>	5.37E-03	5.28E-02
BACU001671	<i>Wdr26</i>	5.17E-05	2.29E-03	BACU014065	<i>CCDC141</i>	5.39E-03	5.28E-02
BACU018425	<i>SH2B3</i>	5.23E-05	2.29E-03	BACU018652	<i>EMILIN3</i>	5.42E-03	5.28E-02
BACU011967	<i>ZNF629</i>	5.35E-05	2.29E-03	BACU004713	<i>ATE1</i>	5.49E-03	5.31E-02
BACU019156	<i>CRIP3</i>	5.93E-05	2.47E-03	BACU014978	<i>TICAM2</i>	5.52E-03	5.31E-02
BACU000562	<i>Tdrd12</i>	6.97E-05	2.84E-03	BACU010396	<i>IFNAR2</i>	5.64E-03	5.35E-02
BACU016544	<i>KNG2</i>	7.48E-05	2.92E-03	BACU002343	<i>ST18</i>	5.64E-03	5.35E-02
BACU019689	<i>NUDT22</i>	7.50E-05	2.92E-03	BACU018699	<i>AAGAB</i>	5.65E-03	5.35E-02
BACU004663	<i>NHLRC2</i>	8.86E-05	3.37E-03	BACU018601	<i>NEK4</i>	5.68E-03	5.35E-02
BACU020364	<i>PRAP1</i>	1.02E-04	3.80E-03	BACU007243	<i>TNFAIP8L2</i>	5.79E-03	5.42E-02
BACU019919	<i>CEP89</i>	1.05E-04	3.83E-03	BACU011206	<i>CXorf58</i>	5.84E-03	5.44E-02
BACU008502	<i>SPTBN5</i>	1.10E-04	3.93E-03	BACU016210	<i>ABL1</i>	6.19E-03	5.74E-02
BACU013216	<i>DERL1</i>	1.18E-04	4.13E-03	BACU005139	<i>TFAP2A</i>	6.27E-03	5.76E-02
BACU000213	<i>C1orf87</i>	1.22E-04	4.16E-03	BACU010470	<i>VPS53</i>	6.28E-03	5.76E-02
BACU019914	<i>CCDC17</i>	1.26E-04	4.16E-03	BACU013909	<i>LRPPRC</i>	6.36E-03	5.80E-02
BACU007951	<i>SCAF11</i>	1.26E-04	4.16E-03	BACU019547	<i>NPPC</i>	6.40E-03	5.81E-02
BACU002375	<i>RASSF1</i>	1.40E-04	4.54E-03	BACU006981	<i>Scrn2</i>	6.47E-03	5.84E-02
BACU005313	<i>BATF2</i>	1.80E-04	5.73E-03	BACU001209	<i>MFSD6L</i>	6.50E-03	5.84E-02
BACU010672	<i>GPR173</i>	1.86E-04	5.82E-03	BACU008948	<i>AMBRA1</i>	6.64E-03	5.93E-02
BACU018780	<i>SLC6A3</i>	1.98E-04	6.09E-03	BACU005484	<i>GPER</i>	6.69E-03	5.93E-02
BACU018825	<i>Cpeb3</i>	2.06E-04	6.22E-03	BACU007600	<i>C10orf52</i>	6.73E-03	5.93E-02
BACU014558	<i>C2orf16</i>	2.37E-04	6.92E-03	BACU009119	<i>Dgkq</i>	6.74E-03	5.93E-02
BACU015937	<i>cystm1</i>	2.37E-04	6.92E-03	BACU007862	<i>TRIM40</i>	6.89E-03	6.04E-02
BACU020473	<i>SLC13A4</i>	3.04E-04	8.70E-03	BACU004407	<i>BACE2</i>	7.02E-03	6.12E-02
BACU001324	<i>SEC24C</i>	3.08E-04	8.70E-03	BACU000320	<i>NASP</i>	7.08E-03	6.14E-02
BACU007890	<i>FAS</i>	3.55E-04	9.87E-03	BACU014478	<i>CCDC82</i>	7.18E-03	6.20E-02
BACU008596	<i>METTL17</i>	3.72E-04	1.02E-02	BACU012683	<i>Xrcc5</i>	7.28E-03	6.25E-02
BACU004705	<i>Rgs10</i>	4.10E-04	1.11E-02	BACU000596	<i>DMKN</i>	7.31E-03	6.25E-02
BACU018892	<i>Atp2b2</i>	4.19E-04	1.11E-02	BACU002405	<i>KLHDC8B</i>	7.36E-03	6.26E-02
BACU014866	<i>C20orf78</i>	4.35E-04	1.14E-02	BACU020078	<i>DTX3L</i>	7.55E-03	6.39E-02
BACU001220	<i>PFA5</i>	4.70E-04	1.20E-02	BACU011656	<i>TBCD</i>	7.88E-03	6.62E-02
BACU006841	<i>SLC13A1</i>	4.77E-04	1.20E-02	BACU010288	<i>BCAP31</i>	7.90E-03	6.62E-02
BACU017483	<i>AKAP12</i>	4.81E-04	1.20E-02	BACU011543	<i>RHBDF2</i>	8.27E-03	6.90E-02
BACU003063	<i>LPCAT3</i>	5.11E-04	1.26E-02	BACU007929	<i>KBTBD12</i>	8.41E-03	6.97E-02
BACU018628	<i>Eml3</i>	5.22E-04	1.27E-02	BACU006751	<i>ZC3HAV1</i>	8.43E-03	6.97E-02
BACU002213	<i>LRRK2</i>	5.44E-04	1.31E-02	BACU015080	<i>Fam3d</i>	8.62E-03	7.09E-02
BACU018595	<i>ANKK</i>	5.66E-04	1.31E-02	BACU000377	<i>ZMYND12</i>	8.76E-03	7.12E-02
BACU019935	<i>LGALS4</i>	5.79E-04	1.31E-02	BACU015164	<i>IL17RC</i>	8.78E-03	7.12E-02
BACU011924	<i>SPN</i>	5.96E-04	1.31E-02	BACU001648	<i>AHCTF1</i>	8.78E-03	7.12E-02

BACU009330	<i>WDR52</i>	6.01E-04	1.31E-02	BACU019836	<i>TMEM168</i>	8.97E-03	7.24E-02
BACU019512	<i>RARS2</i>	6.01E-04	1.31E-02	BACU016289	<i>GOLGA1</i>	9.08E-03	7.30E-02
BACU000158	<i>ZDHHC20</i>	6.02E-04	1.31E-02	BACU013063	<i>IL3</i>	9.23E-03	7.35E-02
BACU011056	<i>CXXC11</i>	6.03E-04	1.31E-02	BACU008463	<i>CFP</i>	9.23E-03	7.35E-02
BACU000666	<i>KIAA1211</i>	6.11E-04	1.31E-02	BACU009285	<i>Iqcb1</i>	9.30E-03	7.37E-02
BACU001908	<i>ECD</i>	6.12E-04	1.31E-02	BACU001839	<i>Ccl22</i>	9.54E-03	7.49E-02
BACU003792	<i>FAM43A</i>	6.50E-04	1.37E-02	BACU009071	<i>TMC4</i>	9.54E-03	7.49E-02
BACU015098	<i>C3orf49</i>	6.65E-04	1.38E-02	BACU019088	<i>AGPAT1</i>	9.72E-03	7.54E-02
BACU002570	<i>Khsrp</i>	6.74E-04	1.38E-02	BACU007776	<i>FMR1NB</i>	9.73E-03	7.54E-02
BACU003818	<i>MUC4</i>	6.76E-04	1.38E-02	BACU001656	<i>C1orf95</i>	9.73E-03	7.54E-02
BACU015246	<i>GATA4</i>	7.16E-04	1.43E-02	BACU002974	<i>AGTRAP</i>	9.93E-03	7.63E-02
BACU012218	<i>Ggnbp1</i>	7.19E-04	1.43E-02	BACU006277	<i>Gm525</i>	9.93E-03	7.63E-02
BACU011156	<i>LY96</i>	7.30E-04	1.44E-02	BACU019926	<i>SEH1L</i>	1.01E-02	7.69E-02
BACU006439	<i>ANKRD40</i>	7.79E-04	1.52E-02	BACU001567	<i>DSG1</i>	1.01E-02	7.69E-02
BACU009991	<i>ALG13</i>	8.02E-04	1.54E-02	BACU013577	<i>ARFIP1</i>	1.02E-02	7.70E-02
BACU001239	<i>Polr2a</i>	8.88E-04	1.68E-02	BACU004795	<i>ZNF511</i>	1.02E-02	7.70E-02
BACU007621	<i>PPYR1</i>	8.93E-04	1.68E-02	BACU015678	<i>ABHD5</i>	1.03E-02	7.71E-02
BACU019917	<i>ODR4</i>	9.34E-04	1.74E-02	BACU006428	<i>TMEM92</i>	1.03E-02	7.71E-02
BACU005661	<i>PDPN</i>	9.47E-04	1.74E-02	BACU000447	<i>VEPH1</i>	1.06E-02	7.87E-02
BACU004035	<i>SLC30A10</i>	9.60E-04	1.74E-02	BACU007560	<i>ADIPOR2</i>	1.06E-02	7.87E-02
BACU015381	<i>DAAMI</i>	9.66E-04	1.74E-02	BACU018176	<i>VAC14</i>	1.07E-02	7.91E-02
BACU008021	<i>LARP4</i>	9.88E-04	1.77E-02	BACU011539	<i>QRICH2</i>	1.11E-02	8.17E-02
BACU004909	<i>CTPS2</i>	1.01E-03	1.79E-02	BACU014502	<i>DCUN1D5</i>	1.12E-02	8.21E-02
BACU006123	<i>Ern1</i>	1.11E-03	1.94E-02	BACU003535	<i>CARNS1</i>	1.14E-02	8.32E-02
BACU002665	<i>SPAG17</i>	1.16E-03	1.99E-02	BACU011209	<i>Prdx4</i>	1.17E-02	8.50E-02
BACU009819	<i>COBRA1</i>	1.16E-03	1.99E-02	BACU007555	<i>DCP1B</i>	1.18E-02	8.54E-02
BACU018235	<i>RABAC1</i>	1.20E-03	2.04E-02	BACU018833	<i>SELPLG</i>	1.20E-02	8.62E-02
BACU002311	<i>NSMAF</i>	1.24E-03	2.06E-02	BACU007516	<i>ZBTB33</i>	1.20E-02	8.62E-02
BACU018055	<i>FANCE</i>	1.26E-03	2.06E-02	BACU014174	<i>BVES</i>	1.21E-02	8.62E-02
BACU009438	<i>CHD1L</i>	1.27E-03	2.06E-02	BACU012082	<i>ZCWPW1</i>	1.21E-02	8.62E-02
BACU006423	<i>SAMD14</i>	1.27E-03	2.06E-02	BACU008888	<i>C11orf41</i>	1.23E-02	8.72E-02
BACU010766	<i>SLC39A9</i>	1.27E-03	2.06E-02	BACU002008	<i>HPSE</i>	1.24E-02	8.76E-02
BACU018158	<i>DHX38</i>	1.30E-03	2.09E-02	BACU009158	<i>CORO7</i>	1.26E-02	8.83E-02
BACU004368	<i>C21orf33</i>	1.34E-03	2.13E-02	BACU002066	<i>CD79A</i>	1.26E-02	8.83E-02
BACU012013	<i>WBSCR28</i>	1.35E-03	2.13E-02	BACU012089	<i>C7orf43</i>	1.27E-02	8.86E-02
BACU017902	<i>FAM219B</i>	1.37E-03	2.14E-02	BACU001695	<i>KIF21B</i>	1.29E-02	8.93E-02
BACU012177	<i>NUFIP1</i>	1.41E-03	2.19E-02	BACU018882	<i>ANKFN1</i>	1.29E-02	8.93E-02
BACU009178	<i>ZSCAN10</i>	1.43E-03	2.20E-02	BACU011903	<i>APOBR</i>	1.30E-02	8.93E-02
BACU003049	<i>BRF2</i>	1.47E-03	2.23E-02	BACU006819	<i>PRRT4</i>	1.30E-02	8.93E-02
BACU009759	<i>CCNB3</i>	1.48E-03	2.23E-02	BACU003261	<i>MED18</i>	1.31E-02	8.93E-02
BACU020226	<i>PTX3</i>	1.49E-03	2.23E-02	BACU003749	<i>SRI</i>	1.31E-02	8.93E-02
BACU016680	<i>GLT25D1</i>	1.50E-03	2.23E-02	BACU010496	<i>DHRS13</i>	1.33E-02	9.03E-02
BACU004797	<i>Caly</i>	1.57E-03	2.31E-02	BACU018958	<i>CAPN7</i>	1.34E-02	9.03E-02
BACU008869	<i>DCDC5</i>	1.64E-03	2.39E-02	BACU001533	<i>TPR</i>	1.34E-02	9.03E-02
BACU013767	<i>NRP1</i>	1.69E-03	2.45E-02	BACU015997	<i>GPR151</i>	1.35E-02	9.06E-02
BACU004080	<i>DDX46</i>	1.75E-03	2.51E-02	BACU014160	<i>ZDBF2</i>	1.37E-02	9.13E-02
BACU011403	<i>Mfsd3</i>	1.92E-03	2.73E-02	BACU002504	<i>Icam4</i>	1.37E-02	9.13E-02
BACU001776	<i>FAM208B</i>	1.95E-03	2.73E-02	BACU015196	<i>TRH</i>	1.38E-02	9.16E-02
BACU009286	<i>GOLGB1</i>	1.95E-03	2.73E-02	BACU006174	<i>WDR41</i>	1.39E-02	9.19E-02
BACU001910	<i>P4HA1</i>	1.97E-03	2.74E-02	BACU013660	<i>TCTN2</i>	1.41E-02	9.25E-02
BACU002680	<i>TSPAN2</i>	2.03E-03	2.77E-02	BACU003753	<i>STK31</i>	1.41E-02	9.25E-02
BACU000375	<i>CCDC30</i>	2.03E-03	2.77E-02	BACU012883	<i>CD24</i>	1.44E-02	9.41E-02
BACU019581	<i>MCMBP</i>	2.04E-03	2.77E-02	BACU018778	<i>SYCE2</i>	1.47E-02	9.57E-02
BACU019545	<i>DECR2</i>	2.07E-03	2.78E-02	BACU013319	<i>NEIL3</i>	1.48E-02	9.60E-02
BACU010459	<i>Rilp</i>	2.08E-03	2.78E-02	BACU019172	<i>Dgat1</i>	1.49E-02	9.63E-02
BACU008003	<i>FAM186B</i>	2.12E-03	2.81E-02	BACU004204	<i>IL2</i>	1.52E-02	9.79E-02
BACU000030	<i>CEP97</i>	2.39E-03	3.15E-02	BACU006861	<i>TOM1L1</i>	1.53E-02	9.82E-02
BACU016786	<i>Klf11</i>	2.42E-03	3.16E-02	BACU012718	<i>CCDC108</i>	1.54E-02	9.85E-02
BACU000230	<i>TMEM61</i>	2.44E-03	3.17E-02	BACU004916	<i>SCML1</i>	1.55E-02	9.87E-02
BACU006752	<i>ZC3HAV1L</i>	2.51E-03	3.23E-02	BACU008714	<i>SRP54</i>	1.56E-02	9.90E-02
BACU017652	<i>Skor2</i>	2.53E-03	3.24E-02	BACU013320	<i>AGA</i>	1.57E-02	9.93E-02
BACU019788	<i>PDE10A</i>	2.73E-03	3.47E-02	BACU020100	<i>MANEA</i>	1.59E-02	9.98E-02
BACU020379	<i>ZACN</i>	2.75E-03	3.47E-02	BACU007748	<i>MTMR1</i>	1.59E-02	9.98E-02
BACU007339	<i>PTGR1</i>	2.77E-03	3.47E-02				

**Supplementary Table 38 | Positively selected genes (PSGs) in the bottlenose dolphin.** The bottlenose dolphin was used as the foreground branch for the PSGs of the bottlenose dolphin. PSGs were selected by a conservative 10% false discovery rate criterion. N/A indicates not available gene name.

Protein ID	Gene Name	P-value	Q-value	Protein ID	Gene Name	P-value	Q-value
ENSTTRP00000014628	<i>SRRM2</i>	3.33E-16	6.06E-13	ENSTTRP00000014126	<i>TMEM196</i>	4.69E-03	3.05E-02
ENSTTRP00000010529	<i>SRRT</i>	5.88E-10	5.35E-07	ENSTTRP00000013856	<i>VEPH1</i>	4.84E-03	3.13E-02
ENSTTRP00000001773	<i>USO1</i>	1.67E-09	7.31E-07	ENSTTRP00000013316	<i>ERO1LB</i>	4.90E-03	3.16E-02
ENSTTRP00000006616	<i>FBNP1L</i>	1.91E-09	7.31E-07	ENSTTRP00000004691	<i>ASTL</i>	4.93E-03	3.17E-02
ENSTTRP00000006238	<i>LENG9</i>	2.01E-09	7.31E-07	ENSTTRP00000003917	<i>C1orf173</i>	5.06E-03	3.24E-02
ENSTTRP00000012835	<i>KIAA1468</i>	3.15E-09	9.55E-07	ENSTTRP00000003634	<i>GFRAL</i>	5.11E-03	3.26E-02
ENSTTRP00000008085	<i>SLC9C2</i>	7.09E-09	1.84E-06	ENSTTRP00000016129	<i>TNFRSF4</i>	5.15E-03	3.27E-02
ENSTTRP00000015712	<i>GNPAT</i>	2.45E-08	5.57E-06	ENSTTRP00000004208	N/A	5.26E-03	3.33E-02
ENSTTRP0000000398	<i>KCNK5</i>	8.12E-08	1.64E-05	ENSTTRP00000008265	<i>MUSK</i>	5.35E-03	3.36E-02
ENSTTRP00000004429	<i>ARHGEF17</i>	9.61E-08	1.75E-05	ENSTTRP00000007734	<i>TMPPRS7</i>	5.35E-03	3.36E-02
ENSTTRP00000007462	<i>TEX11</i>	1.56E-07	2.58E-05	ENSTTRP00000010082	<i>ECM1</i>	5.36E-03	3.36E-02
ENSTTRP00000005095	<i>ZNF518B</i>	2.18E-07	3.26E-05	ENSTTRP00000002847	<i>AAGAB</i>	5.39E-03	3.37E-02
ENSTTRP00000012914	N/A	2.42E-07	3.26E-05	ENSTTRP00000011167	<i>AEBP2</i>	5.49E-03	3.42E-02
ENSTTRP00000006664	<i>TAB2</i>	2.51E-07	3.26E-05	ENSTTRP00000008578	<i>MAOA</i>	5.51E-03	3.42E-02
ENSTTRP00000013530	<i>POLR2A</i>	2.94E-07	3.56E-05	ENSTTRP00000008771	<i>TSNAXIP1</i>	5.61E-03	3.47E-02
ENSTTRP00000010573	<i>CDHR5</i>	6.66E-07	7.33E-05	ENSTTRP00000000860	<i>CHGB</i>	5.65E-03	3.48E-02
ENSTTRP00000001801	<i>PGF</i>	6.85E-07	7.33E-05	ENSTTRP00000012811	<i>C5orf48</i>	5.72E-03	3.51E-02
ENSTTRP00000012251	<i>LRRN4</i>	8.07E-07	7.84E-05	ENSTTRP00000001409	<i>OXA1L</i>	5.82E-03	3.55E-02
ENSTTRP00000007386	<i>OPN4</i>	8.19E-07	7.84E-05	ENSTTRP00000000681	<i>RRP1B</i>	5.82E-03	3.55E-02
ENSTTRP000000011019	<i>PHYHD1</i>	9.82E-07	8.93E-05	ENSTTRP00000001129	<i>C9orf40</i>	5.84E-03	3.55E-02
ENSTTRP00000006118	<i>CYLC1</i>	1.37E-06	1.17E-04	ENSTTRP00000007658	<i>MMAA</i>	5.88E-03	3.56E-02
ENSTTRP00000003737	<i>VSIG10L</i>	1.42E-06	1.17E-04	ENSTTRP00000015646	N/A	5.93E-03	3.58E-02
ENSTTRP00000003373	<i>SYCE2</i>	2.10E-06	1.62E-04	ENSTTRP00000000690	<i>UBASH3A</i>	5.95E-03	3.58E-02
ENSTTRP000000011440	<i>NBR1</i>	2.14E-06	1.62E-04	ENSTTRP00000008936	<i>CCDC57</i>	6.05E-03	3.63E-02
ENSTTRP00000003319	<i>GTPBP8</i>	2.32E-06	1.69E-04	ENSTTRP00000004881	<i>ASNSD1</i>	6.07E-03	3.63E-02
ENSTTRP00000009852	<i>COL3A1</i>	2.88E-06	2.01E-04	ENSTTRP00000011576	N/A	6.29E-03	3.74E-02
ENSTTRP000000016249	<i>VCPIP1</i>	3.90E-06	2.63E-04	ENSTTRP000000011571	<i>OTUD4</i>	6.30E-03	3.74E-02
ENSTTRP00000005149	<i>C19orf43</i>	4.27E-06	2.77E-04	ENSTTRP00000002436	<i>IL5</i>	6.51E-03	3.84E-02
ENSTTRP00000002390	<i>SPTBN5</i>	4.69E-06	2.94E-04	ENSTTRP00000008123	<i>SERPINC1</i>	6.51E-03	3.84E-02
ENSTTRP00000012033	<i>WBSCR27</i>	5.25E-06	3.18E-04	ENSTTRP00000002130	<i>PLAT</i>	6.72E-03	3.96E-02
ENSTTRP00000003696	<i>DCT</i>	6.62E-06	3.88E-04	ENSTTRP00000000085	<i>FAM73A</i>	6.76E-03	3.97E-02
ENSTTRP00000007057	<i>HCFC1</i>	7.04E-06	4.00E-04	ENSTTRP00000007907	<i>ELOVL3</i>	6.79E-03	3.97E-02
ENSTTRP00000000079	<i>HUWE1</i>	8.71E-06	4.71E-04	ENSTTRP00000008942	<i>MUC1</i>	6.85E-03	3.99E-02
ENSTTRP00000001586	<i>CEP89</i>	8.80E-06	4.71E-04	ENSTTRP000000011634	<i>SEH1L</i>	6.90E-03	4.01E-02
ENSTTRP00000014358	<i>DMKN</i>	9.06E-06	4.71E-04	ENSTTRP00000001575	<i>TACO1</i>	7.11E-03	4.12E-02
ENSTTRP000000011043	<i>IGSF21</i>	9.70E-06	4.90E-04	ENSTTRP00000009228	<i>IQCC</i>	7.22E-03	4.17E-02
ENSTTRP00000004453	<i>C1orf87</i>	1.03E-05	5.03E-04	ENSTTRP000000013449	<i>AKAP12</i>	7.28E-03	4.18E-02
ENSTTRP00000013837	<i>RILPL1</i>	1.05E-05	5.03E-04	ENSTTRP00000009200	<i>TRIP4</i>	7.29E-03	4.18E-02
ENSTTRP00000002773	<i>FSTLA</i>	1.10E-05	5.13E-04	ENSTTRP00000001636	<i>ITGAL</i>	7.42E-03	4.24E-02
ENSTTRP00000001281	<i>C12orf29</i>	1.23E-05	5.59E-04	ENSTTRP00000003146	<i>CHADL</i>	7.53E-03	4.29E-02
ENSTTRP00000008528	<i>ILIRAP</i>	1.75E-05	7.76E-04	ENSTTRP000000015700	<i>VGLL4</i>	7.64E-03	4.33E-02
ENSTTRP00000005779	<i>NOSTRIN</i>	1.81E-05	7.84E-04	ENSTTRP00000000837	<i>NUDT15</i>	7.65E-03	4.33E-02
ENSTTRP00000011073	<i>ZNF750</i>	1.86E-05	7.87E-04	ENSTTRP000000014129	<i>CCDC63</i>	7.84E-03	4.43E-02
ENSTTRP00000003880	<i>FBXL7</i>	1.92E-05	7.94E-04	ENSTTRP00000002597	<i>MDGA2</i>	7.89E-03	4.43E-02
ENSTTRP00000013717	<i>TAPBP1L</i>	2.48E-05	1.00E-03	ENSTTRP00000016218	<i>SLC38A1</i>	7.89E-03	4.43E-02
ENSTTRP00000009637	<i>ADAMTS20</i>	2.62E-05	1.04E-03	ENSTTRP000000006711	<i>LRRC6</i>	7.94E-03	4.44E-02
ENSTTRP00000000301	<i>HIBCH</i>	2.76E-05	1.07E-03	ENSTTRP000000005299	<i>SLC16A8</i>	7.97E-03	4.45E-02
ENSTTRP00000007943	<i>ENAM</i>	2.83E-05	1.07E-03	ENSTTRP00000004260	<i>KIAA0922</i>	8.11E-03	4.51E-02
ENSTTRP00000006884	<i>CMPK2</i>	3.17E-05	1.18E-03	ENSTTRP00000015438	<i>LRCH3</i>	8.17E-03	4.53E-02
ENSTTRP00000010352	<i>GABRP</i>	3.40E-05	1.24E-03	ENSTTRP00000008272	<i>IL10RA</i>	8.27E-03	4.56E-02
ENSTTRP000000013913	<i>C12orf29</i>	3.75E-05	1.33E-03	ENSTTRP00000005063	<i>KIAA1191</i>	8.30E-03	4.56E-02
ENSTTRP00000010370	<i>MEP1B</i>	3.80E-05	1.33E-03	ENSTTRP00000007783	<i>MIEN1</i>	8.32E-03	4.56E-02
ENSTTRP00000009850	<i>KHSRP</i>	4.48E-05	1.54E-03	ENSTTRP00000009885	<i>C5</i>	8.33E-03	4.56E-02
ENSTTRP00000004377	<i>GLO1</i>	4.69E-05	1.56E-03	ENSTTRP000000009724	<i>SMTN</i>	8.35E-03	4.56E-02
ENSTTRP00000015665	<i>AK3</i>	4.72E-05	1.56E-03	ENSTTRP00000006308	<i>KAT2A</i>	8.40E-03	4.56E-02
ENSTTRP00000015855	<i>CCDC116</i>	5.22E-05	1.70E-03	ENSTTRP00000003020	<i>DOT1L</i>	8.41E-03	4.56E-02
ENSTTRP00000008565	<i>DPM1</i>	5.90E-05	1.88E-03	ENSTTRP000000008641	<i>IL17RD</i>	8.44E-03	4.56E-02
ENSTTRP00000006080	N/A	7.43E-05	2.32E-03	ENSTTRP000000000618	<i>VSIG10</i>	8.45E-03	4.56E-02
ENSTTRP00000008080	<i>PIGB</i>	7.53E-05	2.32E-03	ENSTTRP00000003677	<i>NAP1L5</i>	8.55E-03	4.60E-02
ENSTTRP00000015059	<i>TRIM47</i>	8.03E-05	2.43E-03	ENSTTRP00000003165	<i>KBTBD3</i>	8.71E-03	4.67E-02
ENSTTRP00000002044	<i>LCP2</i>	8.43E-05	2.51E-03	ENSTTRP00000005670	<i>C22orf13</i>	8.85E-03	4.73E-02
ENSTTRP00000010375	<i>MMS22L</i>	9.89E-05	2.83E-03	ENSTTRP00000002816	<i>ARPC1A</i>	8.89E-03	4.74E-02
ENSTTRP00000000582	<i>CATSPER4</i>	9.93E-05	2.83E-03	ENSTTRP000000011581	<i>TDRD12</i>	9.02E-03	4.77E-02
ENSTTRP00000007283	<i>SLC47A2</i>	1.00E-04	2.83E-03	ENSTTRP00000004493	<i>MLL2</i>	9.02E-03	4.77E-02
ENSTTRP000000016071	N/A	1.01E-04	2.83E-03	ENSTTRP00000002534	<i>ANLN</i>	9.03E-03	4.77E-02
ENSTTRP00000007863	<i>RABAC1</i>	1.03E-04	2.84E-03	ENSTTRP00000002369	<i>CD38</i>	9.08E-03	4.79E-02
ENSTTRP00000005938	<i>ACBD4</i>	1.14E-04	3.09E-03	ENSTTRP00000004043	<i>TTC36</i>	9.23E-03	4.85E-02
ENSTTRP00000003607	<i>LIMCH1</i>	1.16E-04	3.10E-03	ENSTTRP000000010201	<i>JMJD4</i>	9.35E-03	4.90E-02
ENSTTRP00000011565	<i>CCDC17</i>	1.25E-04	3.29E-03	ENSTTRP00000016378	<i>ENTHD2</i>	9.50E-03	4.95E-02
ENSTTRP00000010378	<i>KCNJ13</i>	1.41E-04	3.61E-03	ENSTTRP00000016089	<i>CCDC22</i>	9.50E-03	4.95E-02
ENSTTRP00000002487	<i>RABL6</i>	1.42E-04	3.61E-03	ENSTTRP00000003398	<i>SLC6A3</i>	9.54E-03	4.96E-02

ENSTTRP00000012100	<i>BIN3</i>	1.43E-04	3.61E-03	ENSTTRP00000004221	<i>C2orf73</i>	9.57E-03	4.96E-02
ENSTTRP00000003683	<i>LRFN3</i>	1.54E-04	3.84E-03	ENSTTRP00000016298	N/A	9.64E-03	4.97E-02
ENSTTRP000000014256	<i>MARCKS</i>	1.63E-04	4.00E-03	ENSTTRP00000000830	<i>PANX2</i>	9.66E-03	4.97E-02
ENSTTRP00000003447	<i>ENGASE</i>	1.65E-04	4.00E-03	ENSTTRP00000011113	<i>C9</i>	9.67E-03	4.97E-02
ENSTTRP00000009854	<i>PPP1R12A</i>	1.67E-04	4.00E-03	ENSTTRP00000007761	<i>MLST8</i>	9.82E-03	5.03E-02
ENSTTRP00000009059	<i>XPO6</i>	1.70E-04	4.02E-03	ENSTTRP00000005974	<i>PARP10</i>	9.93E-03	5.07E-02
ENSTTRP00000012386	<i>NDUFAF6</i>	1.82E-04	4.24E-03	ENSTTRP00000006539	<i>ETV3L</i>	1.01E-02	5.13E-02
ENSTTRP00000014694	<i>NPHP3</i>	1.91E-04	4.36E-03	ENSTTRP00000015668	<i>IL33</i>	1.01E-02	5.13E-02
ENSTTRP00000002342	<i>FAM129C</i>	1.92E-04	4.36E-03	ENSTTRP00000015277	<i>PITHD1</i>	1.02E-02	5.17E-02
ENSTTRP000000015515	<i>CD59</i>	2.12E-04	4.76E-03	ENSTTRP00000006620	<i>MLLT11</i>	1.03E-02	5.20E-02
ENSTTRP00000009536	<i>ANXA1</i>	2.30E-04	5.10E-03	ENSTTRP00000005536	<i>STRC</i>	1.04E-02	5.22E-02
ENSTTRP00000007924	<i>FRA10AC1</i>	2.35E-04	5.15E-03	ENSTTRP00000007688	N/A	1.04E-02	5.22E-02
ENSTTRP00000004433	<i>SON</i>	2.42E-04	5.24E-03	ENSTTRP00000006119	<i>FAIM</i>	1.05E-02	5.25E-02
ENSTTRP00000002178	<i>MUC13</i>	2.54E-04	5.43E-03	ENSTTRP00000001377	<i>DCTD</i>	1.05E-02	5.25E-02
ENSTTRP000000016116	<i>OLA1</i>	2.57E-04	5.43E-03	ENSTTRP00000015611	<i>ST6GALNAC6</i>	1.07E-02	5.32E-02
ENSTTRP00000009041	<i>ICAM4</i>	2.60E-04	5.44E-03	ENSTTRP00000013197	<i>CNFN</i>	1.07E-02	5.32E-02
ENSTTRP000000011307	<i>ARHGEF39</i>	2.76E-04	5.68E-03	ENSTTRP00000013260	<i>LEMD1</i>	1.08E-02	5.35E-02
ENSTTRP00000012681	<i>AHSA2</i>	2.78E-04	5.68E-03	ENSTTRP00000007849	N/A	1.09E-02	5.39E-02
ENSTTRP00000014958	<i>SCFD2</i>	3.01E-04	6.06E-03	ENSTTRP00000002191	<i>SPINK5</i>	1.10E-02	5.42E-02
ENSTTRP00000004673	<i>DCLK3</i>	3.03E-04	6.06E-03	ENSTTRP00000005023	<i>CD37</i>	1.11E-02	5.46E-02
ENSTTRP00000006398	<i>CXorf36</i>	3.19E-04	6.31E-03	ENSTTRP00000011292	<i>INSC</i>	1.12E-02	5.49E-02
ENSTTRP00000014366	<i>KLRG2</i>	3.26E-04	6.37E-03	ENSTTRP00000002741	<i>FIBIN</i>	1.13E-02	5.52E-02
ENSTTRP00000003515	<i>AP4S1</i>	3.29E-04	6.37E-03	ENSTTRP00000013691	<i>RPAP2</i>	1.14E-02	5.56E-02
ENSTTRP00000012543	<i>CXorf26</i>	3.63E-04	6.88E-03	ENSTTRP00000009708	<i>FBXL8</i>	1.15E-02	5.59E-02
ENSTTRP00000000211	<i>SMARCAD1</i>	3.63E-04	6.88E-03	ENSTTRP00000009233	<i>CLCN3</i>	1.16E-02	5.63E-02
ENSTTRP00000007588	<i>FAS</i>	3.82E-04	7.16E-03	ENSTTRP00000009672	<i>C11orf82</i>	1.17E-02	5.66E-02
ENSTTRP00000003918	<i>SOS2</i>	4.03E-04	7.48E-03	ENSTTRP00000012728	<i>IL17RA</i>	1.18E-02	5.68E-02
ENSTTRP00000008154	<i>PTCD3</i>	4.27E-04	7.84E-03	ENSTTRP00000005513	<i>WDR52</i>	1.18E-02	5.68E-02
ENSTTRP00000004135	<i>METAP2</i>	4.31E-04	7.84E-03	ENSTTRP00000000625	<i>FAM19A4</i>	1.19E-02	5.68E-02
ENSTTRP000000014169	<i>C19orf26</i>	4.43E-04	7.98E-03	ENSTTRP00000002201	<i>EIF2AK3</i>	1.19E-02	5.68E-02
ENSTTRP00000011474	<i>GRIK4</i>	4.62E-04	8.24E-03	ENSTTRP00000011347	<i>AVP</i>	1.19E-02	5.68E-02
ENSTTRP00000013874	<i>NARS2</i>	4.74E-04	8.37E-03	ENSTTRP00000001080	<i>SCARB1</i>	1.20E-02	5.71E-02
ENSTTRP00000004296	<i>INTS2</i>	4.85E-04	8.47E-03	ENSTTRP00000008006	<i>AHCYL2</i>	1.21E-02	5.75E-02
ENSTTRP000000015575	<i>C20orf72</i>	4.89E-04	8.47E-03	ENSTTRP00000012464	<i>FBRS</i>	1.22E-02	5.78E-02
ENSTTRP00000000906	<i>PDC12</i>	5.00E-04	8.58E-03	ENSTTRP00000010630	<i>TNFRSF12A</i>	1.23E-02	5.81E-02
ENSTTRP00000004316	<i>GTPBP10</i>	5.05E-04	8.58E-03	ENSTTRP00000002651	N/A	1.24E-02	5.81E-02
ENSTTRP000000010920	<i>GOLGB1</i>	5.28E-04	8.89E-03	ENSTTRP00000013146	<i>C8orf33</i>	1.24E-02	5.81E-02
ENSTTRP00000002931	<i>MYO9B</i>	5.38E-04	8.93E-03	ENSTTRP00000009446	<i>NRP1</i>	1.24E-02	5.81E-02
ENSTTRP00000009461	<i>WEE2</i>	5.40E-04	8.93E-03	ENSTTRP00000003594	<i>GLOD4</i>	1.25E-02	5.83E-02
ENSTTRP00000007320	<i>KDM1A</i>	5.61E-04	9.16E-03	ENSTTRP00000009531	<i>CX3CL1</i>	1.25E-02	5.83E-02
ENSTTRP00000000056	<i>PML</i>	5.64E-04	9.16E-03	ENSTTRP00000012204	<i>PTPN21</i>	1.26E-02	5.86E-02
ENSTTRP00000003107	<i>UBE2N</i>	5.76E-04	9.27E-03	ENSTTRP00000010895	N/A	1.27E-02	5.89E-02
ENSTTRP00000013898	<i>C18orf54</i>	5.87E-04	9.33E-03	ENSTTRP00000014372	<i>METTL14</i>	1.29E-02	5.97E-02
ENSTTRP000000008694	<i>DECR2</i>	5.90E-04	9.33E-03	ENSTTRP00000009486	<i>CFP</i>	1.30E-02	6.00E-02
ENSTTRP00000002442	<i>KIF3A</i>	6.21E-04	9.74E-03	ENSTTRP00000010555	<i>ATP10A</i>	1.31E-02	6.03E-02
ENSTTRP00000015468	<i>C11orf58</i>	6.28E-04	9.76E-03	ENSTTRP00000013761	<i>SH2D2A</i>	1.32E-02	6.06E-02
ENSTTRP000000008741	<i>APOA2</i>	6.47E-04	9.97E-03	ENSTTRP00000013325	N/A	1.33E-02	6.08E-02
ENSTTRP00000000026	N/A	6.62E-04	1.00E-02	ENSTTRP00000007773	<i>E4F1</i>	1.33E-02	6.08E-02
ENSTTRP00000016206	<i>HRC</i>	6.63E-04	1.00E-02	ENSTTRP00000014202	<i>NGLY1</i>	1.35E-02	6.15E-02
ENSTTRP00000003046	N/A	6.79E-04	1.02E-02	ENSTTRP00000007445	<i>KCNJ14</i>	1.36E-02	6.18E-02
ENSTTRP00000009682	<i>TXNDC2</i>	7.26E-04	1.06E-02	ENSTTRP00000007506	<i>DUSP15</i>	1.37E-02	6.20E-02
ENSTTRP00000011662	N/A	7.30E-04	1.06E-02	ENSTTRP00000001191	N/A	1.37E-02	6.20E-02
ENSTTRP00000014950	<i>C6orf170</i>	7.31E-04	1.06E-02	ENSTTRP00000000854	N/A	1.38E-02	6.23E-02
ENSTTRP000000006636	<i>SRF</i>	7.34E-04	1.06E-02	ENSTTRP00000013252	<i>NCAPH2</i>	1.39E-02	6.26E-02
ENSTTRP00000013813	<i>LONRF1</i>	7.35E-04	1.06E-02	ENSTTRP00000008413	<i>KDM5A</i>	1.40E-02	6.26E-02
ENSTTRP00000014484	<i>C18orf63</i>	7.39E-04	1.06E-02	ENSTTRP00000015959	<i>FETUB</i>	1.40E-02	6.26E-02
ENSTTRP00000015490	<i>TMIGD1</i>	7.63E-04	1.08E-02	ENSTTRP00000005990	<i>VSTM4</i>	1.40E-02	6.26E-02
ENSTTRP000000006507	<i>EIF2AK2</i>	7.73E-04	1.09E-02	ENSTTRP00000009518	<i>ANKZF1</i>	1.41E-02	6.27E-02
ENSTTRP00000009277	<i>P4HA1</i>	7.77E-04	1.09E-02	ENSTTRP00000008104	<i>ZFP57</i>	1.41E-02	6.27E-02
ENSTTRP00000010106	<i>CHFR</i>	7.84E-04	1.09E-02	ENSTTRP00000006921	<i>TLL6</i>	1.42E-02	6.30E-02
ENSTTRP00000006504	<i>CCDC75</i>	8.12E-04	1.12E-02	ENSTTRP00000015382	<i>ING3</i>	1.43E-02	6.31E-02
ENSTTRP00000005160	<i>ASNA1</i>	8.27E-04	1.13E-02	ENSTTRP00000014723	<i>RP1L1</i>	1.43E-02	6.31E-02
ENSTTRP00000003486	<i>C21orf56</i>	8.42E-04	1.14E-02	ENSTTRP00000009335	<i>FASTKD2</i>	1.44E-02	6.34E-02
ENSTTRP000000001622	<i>PLIN4</i>	8.54E-04	1.15E-02	ENSTTRP00000010391	<i>GUSB</i>	1.45E-02	6.34E-02
ENSTTRP00000007953	<i>CTSZ</i>	8.96E-04	1.20E-02	ENSTTRP00000010442	<i>C21orf91</i>	1.46E-02	6.34E-02
ENSTTRP00000001755	N/A	9.09E-04	1.21E-02	ENSTTRP00000005975	<i>RAB18</i>	1.46E-02	6.34E-02
ENSTTRP000000004443	<i>HDAC6</i>	9.45E-04	1.23E-02	ENSTTRP00000006135	<i>LRR1Q3</i>	1.46E-02	6.34E-02
ENSTTRP000000013933	<i>CCDC39</i>	9.47E-04	1.23E-02	ENSTTRP00000007923	<i>C19orf57</i>	1.47E-02	6.34E-02
ENSTTRP00000000612	<i>GUCY1A3</i>	9.54E-04	1.23E-02	ENSTTRP00000011490	<i>FCAMR</i>	1.47E-02	6.34E-02
ENSTTRP00000000330	<i>CCDC129</i>	9.58E-04	1.23E-02	ENSTTRP00000013046	<i>HPS5</i>	1.47E-02	6.34E-02
ENSTTRP000000010518	<i>PDE10A</i>	9.62E-04	1.23E-02	ENSTTRP00000004641	<i>MITD1</i>	1.47E-02	6.34E-02
ENSTTRP00000000105	<i>PIK3C3</i>	9.64E-04	1.23E-02	ENSTTRP00000001740	<i>MED18</i>	1.47E-02	6.34E-02
ENSTTRP00000014876	<i>BHLHA15</i>	9.78E-04	1.24E-02	ENSTTRP00000013370	<i>BAG6</i>	1.49E-02	6.41E-02
ENSTTRP00000005475	<i>BOC</i>	9.89E-04	1.24E-02	ENSTTRP00000011346	<i>SH3D21</i>	1.50E-02	6.43E-02
ENSTTRP00000013436	<i>MTCH2</i>	1.00E-03	1.24E-02	ENSTTRP00000009163	<i>IQCE</i>	1.51E-02	6.45E-02
ENSTTRP00000002952	<i>TDRKH</i>	1.01E-03	1.24E-02	ENSTTRP00000010775	<i>AMACR</i>	1.51E-02	6.45E-02
ENSTTRP00000015087	<i>QRICH2</i>	1.02E-03	1.24E-02	ENSTTRP000000008740	<i>RAPGEF1</i>	1.55E-02	6.59E-02
ENSTTRP00000003289	<i>AP5M1</i>	1.02E-03	1.24E-02	ENSTTRP00000001173	<i>PPP1R15A</i>	1.55E-02	6.59E-02
ENSTTRP00000001970	<i>MED19</i>	1.03E-03	1.25E-02	ENSTTRP00000007313	<i>NFATC2</i>	1.56E-02	6.60E-02
ENSTTRP00000005481	<i>UPRT</i>	1.04E-03	1.25E-02	ENSTTRP00000014728	<i>PRSS53</i>	1.56E-02	6.60E-02

ENSTTRP00000008520	<i>BEND2</i>	1.14E-03	1.36E-02	ENSTTRP00000001806	<i>MLH3</i>	1.58E-02	6.67E-02
ENSTTRP00000003840	<i>C1orf112</i>	1.15E-03	1.37E-02	ENSTTRP00000008529	<i>GA56</i>	1.60E-02	6.74E-02
ENSTTRP00000008294	<i>ASB2</i>	1.16E-03	1.37E-02	ENSTTRP00000012472	<i>HEXDC</i>	1.61E-02	6.75E-02
ENSTTRP00000000039	<i>C15orf52</i>	1.18E-03	1.38E-02	ENSTTRP00000007264	<i>REXO4</i>	1.61E-02	6.75E-02
ENSTTRP00000015812	<i>NR0B2</i>	1.20E-03	1.40E-02	ENSTTRP00000003524	<i>SPATA24</i>	1.62E-02	6.76E-02
ENSTTRP00000001993	<i>C10orf47</i>	1.23E-03	1.42E-02	ENSTTRP00000005731	<i>NPAT</i>	1.62E-02	6.76E-02
ENSTTRP00000001154	<i>PDPN</i>	1.23E-03	1.42E-02	ENSTTRP00000012568	<i>TDRD5</i>	1.64E-02	6.79E-02
ENSTTRP00000001023	<i>CSF2RB</i>	1.26E-03	1.44E-02	ENSTTRP00000007140	<i>IL17RE</i>	1.64E-02	6.79E-02
ENSTTRP00000005349	<i>SCML1</i>	1.31E-03	1.49E-02	ENSTTRP00000012559	<i>AGA</i>	1.64E-02	6.79E-02
ENSTTRP00000001238	<i>ZNRF3</i>	1.32E-03	1.49E-02	ENSTTRP00000006832	<i>KCNA6</i>	1.65E-02	6.82E-02
ENSTTRP00000001188	<i>CBX7</i>	1.34E-03	1.50E-02	ENSTTRP00000004537	<i>LPCAT3</i>	1.66E-02	6.83E-02
ENSTTRP00000008502	<i>CAPG</i>	1.44E-03	1.60E-02	ENSTTRP00000002248	<i>LRRK2</i>	1.66E-02	6.83E-02
ENSTTRP00000000640	<i>MX2</i>	1.44E-03	1.60E-02	ENSTTRP00000001150	<i>LRRC38</i>	1.67E-02	6.84E-02
ENSTTRP00000007673	<i>LEF1</i>	1.45E-03	1.60E-02	ENSTTRP00000010735	<i>FAM129A</i>	1.67E-02	6.84E-02
ENSTTRP00000013862	<i>TNPO3</i>	1.48E-03	1.62E-02	ENSTTRP00000007545	<i>SEMA7A</i>	1.68E-02	6.84E-02
ENSTTRP00000010663	<i>ALG3</i>	1.54E-03	1.68E-02	ENSTTRP00000010280	<i>POLR2B</i>	1.68E-02	6.84E-02
ENSTTRP00000010840	<i>RIBC1</i>	1.58E-03	1.70E-02	ENSTTRP00000015327	<i>DCST2</i>	1.68E-02	6.84E-02
ENSTTRP00000002820	N/A	1.59E-03	1.70E-02	ENSTTRP00000005110	<i>KIF12</i>	1.69E-02	6.86E-02
ENSTTRP00000001337	<i>CLN5</i>	1.60E-03	1.70E-02	ENSTTRP00000007345	<i>FOS</i>	1.70E-02	6.89E-02
ENSTTRP00000009008	<i>INTS6</i>	1.60E-03	1.70E-02	ENSTTRP00000005254	<i>FOXP2</i>	1.71E-02	6.91E-02
ENSTTRP00000007122	<i>CAPS</i>	1.63E-03	1.71E-02	ENSTTRP00000000417	<i>AMBRA1</i>	1.73E-02	6.98E-02
ENSTTRP00000012156	<i>KIAA0284</i>	1.63E-03	1.71E-02	ENSTTRP00000010921	<i>IQCB1</i>	1.74E-02	6.99E-02
ENSTTRP00000008312	<i>C19orf68</i>	1.66E-03	1.74E-02	ENSTTRP00000005476	<i>LRRC27</i>	1.74E-02	6.99E-02
ENSTTRP00000012271	<i>SH3PXD2B</i>	1.69E-03	1.76E-02	ENSTTRP00000003349	<i>SUPT5H</i>	1.77E-02	7.09E-02
ENSTTRP00000000494	<i>RNF213</i>	1.71E-03	1.77E-02	ENSTTRP00000002394	<i>IL20RA</i>	1.78E-02	7.11E-02
ENSTTRP00000004699	<i>CDKL5</i>	1.77E-03	1.82E-02	ENSTTRP00000009793	<i>MTERF</i>	1.79E-02	7.12E-02
ENSTTRP000000011460	<i>KIAA0913</i>	1.79E-03	1.82E-02	ENSTTRP00000000291	<i>NOTCH2</i>	1.79E-02	7.12E-02
ENSTTRP00000015787	<i>SVIL</i>	1.79E-03	1.82E-02	ENSTTRP00000004149	<i>SPATA2L</i>	1.81E-02	7.13E-02
ENSTTRP00000009240	<i>POLK</i>	1.81E-03	1.83E-02	ENSTTRP00000005231	<i>TSKS</i>	1.81E-02	7.13E-02
ENSTTRP00000010762	<i>RSPH3</i>	1.82E-03	1.83E-02	ENSTTRP00000010892	<i>TNRC6B</i>	1.81E-02	7.13E-02
ENSTTRP00000010472	<i>CKAP2L</i>	1.85E-03	1.84E-02	ENSTTRP00000007855	<i>C2orf71</i>	1.81E-02	7.13E-02
ENSTTRP00000007759	<i>SLC9C1</i>	1.85E-03	1.84E-02	ENSTTRP00000008001	<i>TDRD6</i>	1.81E-02	7.13E-02
ENSTTRP00000016063	<i>BRF2</i>	1.88E-03	1.85E-02	ENSTTRP00000002895	<i>NPHP1</i>	1.82E-02	7.15E-02
ENSTTRP00000007730	<i>SRGN</i>	1.88E-03	1.85E-02	ENSTTRP00000010367	<i>CARD6</i>	1.84E-02	7.21E-02
ENSTTRP00000002690	<i>TRAPPC9</i>	1.90E-03	1.86E-02	ENSTTRP00000002549	<i>ASUN</i>	1.85E-02	7.23E-02
ENSTTRP00000010649	<i>TMEM52</i>	1.95E-03	1.90E-02	ENSTTRP00000007995	<i>GC</i>	1.86E-02	7.23E-02
ENSTTRP000000015944	<i>LRF1</i>	1.96E-03	1.90E-02	ENSTTRP00000009881	<i>AIRE</i>	1.86E-02	7.23E-02
ENSTTRP00000001323	<i>IRAK1BP1</i>	1.97E-03	1.90E-02	ENSTTRP00000013644	<i>C19orf25</i>	1.86E-02	7.23E-02
ENSTTRP00000014985	<i>PRAP1</i>	1.98E-03	1.90E-02	ENSTTRP00000016101	<i>CYP46A1</i>	1.90E-02	7.37E-02
ENSTTRP00000001214	<i>CCDC24</i>	2.17E-03	2.04E-02	ENSTTRP00000013803	<i>MAMSTR</i>	1.91E-02	7.38E-02
ENSTTRP00000012859	<i>PPYR1</i>	2.18E-03	2.04E-02	ENSTTRP00000006654	<i>UGP2</i>	1.91E-02	7.38E-02
ENSTTRP00000011637	<i>CCDC81</i>	2.18E-03	2.04E-02	ENSTTRP00000005057	<i>GOLGA4</i>	1.92E-02	7.39E-02
ENSTTRP00000008585	<i>CD44</i>	2.18E-03	2.04E-02	ENSTTRP00000002840	<i>NSUN7</i>	1.93E-02	7.39E-02
ENSTTRP00000007952	<i>ABHD11</i>	2.21E-03	2.06E-02	ENSTTRP00000002010	<i>CCDC173</i>	1.93E-02	7.39E-02
ENSTTRP00000000206	<i>ANKRD40</i>	2.25E-03	2.09E-02	ENSTTRP00000012356	<i>NPC2</i>	1.93E-02	7.39E-02
ENSTTRP00000007998	<i>SLC26A2</i>	2.26E-03	2.09E-02	ENSTTRP00000001172	<i>PLEKHA4</i>	1.95E-02	7.45E-02
ENSTTRP000000011134	<i>RUNC1</i>	2.27E-03	2.09E-02	ENSTTRP00000001600	<i>CCDC94</i>	2.00E-02	7.63E-02
ENSTTRP00000007723	<i>BRCC3</i>	2.33E-03	2.13E-02	ENSTTRP00000014155	<i>FBXW4</i>	2.01E-02	7.65E-02
ENSTTRP00000000577	<i>FMR1NB</i>	2.38E-03	2.16E-02	ENSTTRP00000011768	<i>TMEM232</i>	2.02E-02	7.65E-02
ENSTTRP00000010588	<i>IL18RAP</i>	2.40E-03	2.17E-02	ENSTTRP00000009548	<i>CCDC135</i>	2.02E-02	7.65E-02
ENSTTRP00000008519	<i>XCLL1</i>	2.42E-03	2.18E-02	ENSTTRP00000010088	<i>RIOK2</i>	2.06E-02	7.79E-02
ENSTTRP00000001989	<i>CCDC73</i>	2.51E-03	2.24E-02	ENSTTRP00000001570	<i>TANC2</i>	2.08E-02	7.85E-02
ENSTTRP00000008952	<i>TCN1</i>	2.51E-03	2.24E-02	ENSTTRP00000002509	<i>RARRS2</i>	2.09E-02	7.87E-02
ENSTTRP00000007790	<i>GSDMB</i>	2.59E-03	2.29E-02	ENSTTRP00000001703	<i>VSIG8</i>	2.10E-02	7.89E-02
ENSTTRP00000012256	<i>UBA2</i>	2.59E-03	2.29E-02	ENSTTRP00000004469	<i>TP11</i>	2.12E-02	7.95E-02
ENSTTRP00000001503	<i>LEP</i>	2.63E-03	2.29E-02	ENSTTRP00000014145	<i>PTOVI</i>	2.13E-02	7.97E-02
ENSTTRP00000012161	<i>WT1</i>	2.64E-03	2.29E-02	ENSTTRP00000006075	N/A	2.14E-02	7.98E-02
ENSTTRP00000004279	N/A	2.65E-03	2.29E-02	ENSTTRP00000010990	<i>CD274</i>	2.14E-02	7.98E-02
ENSTTRP00000013496	<i>SLC2A6</i>	2.65E-03	2.29E-02	ENSTTRP00000005659	<i>RGSI</i>	2.15E-02	8.00E-02
ENSTTRP00000014092	<i>URB2</i>	2.67E-03	2.30E-02	ENSTTRP00000000095	<i>TEX15</i>	2.17E-02	8.05E-02
ENSTTRP00000006749	<i>SLC13A3</i>	2.68E-03	2.30E-02	ENSTTRP00000008790	<i>IQCJ</i>	2.18E-02	8.07E-02
ENSTTRP00000006629	<i>MRPL2</i>	2.76E-03	2.36E-02	ENSTTRP00000004654	<i>LYG2</i>	2.19E-02	8.10E-02
ENSTTRP00000013482	N/A	2.82E-03	2.38E-02	ENSTTRP00000004721	<i>MESDC1</i>	2.22E-02	8.19E-02
ENSTTRP000000015471	<i>PIK3C2A</i>	2.83E-03	2.38E-02	ENSTTRP00000005614	<i>DOM3Z</i>	2.23E-02	8.21E-02
ENSTTRP00000007270	<i>ADAMTS13</i>	2.83E-03	2.38E-02	ENSTTRP00000008992	<i>FMO1</i>	2.24E-02	8.23E-02
ENSTTRP00000000569	<i>MYBBP1A</i>	2.86E-03	2.38E-02	ENSTTRP00000003083	<i>MBP</i>	2.25E-02	8.24E-02
ENSTTRP00000007675	<i>GALC</i>	2.87E-03	2.38E-02	ENSTTRP00000009781	<i>CNTROB</i>	2.26E-02	8.24E-02
ENSTTRP000000007301	<i>TTC3</i>	2.89E-03	2.38E-02	ENSTTRP00000011851	<i>TMEM132A</i>	2.26E-02	8.24E-02
ENSTTRP00000002373	<i>CLSTN2</i>	2.89E-03	2.38E-02	ENSTTRP00000010639	<i>SNAPC3</i>	2.26E-02	8.24E-02
ENSTTRP00000008049	<i>GLTSCR1</i>	2.89E-03	2.38E-02	ENSTTRP00000009477	<i>ZNF365</i>	2.27E-02	8.26E-02
ENSTTRP000000010316	<i>SCIMP</i>	2.92E-03	2.39E-02	ENSTTRP00000006579	<i>GHR</i>	2.29E-02	8.30E-02
ENSTTRP00000006745	<i>MEGF9</i>	2.93E-03	2.39E-02	ENSTTRP00000011139	<i>BRCA1</i>	2.30E-02	8.30E-02
ENSTTRP000000004164	<i>FANCA</i>	2.95E-03	2.40E-02	ENSTTRP00000008902	<i>ZNF831</i>	2.30E-02	8.30E-02
ENSTTRP00000012238	<i>GRM6</i>	2.99E-03	2.42E-02	ENSTTRP00000012798	<i>PARP14</i>	2.30E-02	8.30E-02
ENSTTRP00000006820	<i>EPB41L3</i>	3.02E-03	2.43E-02	ENSTTRP00000000087	<i>NEXN</i>	2.32E-02	8.34E-02
ENSTTRP00000007099	<i>THEMIS2</i>	3.05E-03	2.43E-02	ENSTTRP00000009547	<i>GPR97</i>	2.32E-02	8.34E-02
ENSTTRP00000016064	<i>RAB11FIP1</i>	3.06E-03	2.43E-02	ENSTTRP00000014813	<i>C4BPA</i>	2.39E-02	8.57E-02
ENSTTRP00000001823	<i>VPS28</i>	3.06E-03	2.43E-02	ENSTTRP00000008831	<i>TMEM237</i>	2.40E-02	8.59E-02
ENSTTRP00000011377	<i>IFT140</i>	3.12E-03	2.47E-02	ENSTTRP00000010871	<i>MCCC2</i>	2.43E-02	8.68E-02
ENSTTRP00000000623	<i>GRIN3A</i>	3.14E-03	2.47E-02	ENSTTRP00000009815	<i>MRII</i>	2.44E-02	8.70E-02

ENSTTRP00000011638	<i>ME3</i>	3.15E-03	2.47E-02	ENSTTRP00000010431	<i>IFRD1</i>	2.45E-02	8.72E-02
ENSTTRP0000000593	<i>PUS1</i>	3.17E-03	2.47E-02	ENSTTRP0000001768	<i>LRRC34</i>	2.48E-02	8.81E-02
ENSTTRP00000005913	<i>SGOL2</i>	3.19E-03	2.48E-02	ENSTTRP00000003136	<i>EP300</i>	2.49E-02	8.83E-02
ENSTTRP00000007203	<i>LMBRD1</i>	3.32E-03	2.57E-02	ENSTTRP00000011621	<i>SFTA2</i>	2.50E-02	8.83E-02
ENSTTRP00000012721	<i>DTX3L</i>	3.35E-03	2.57E-02	ENSTTRP00000011022	<i>FAM73B</i>	2.50E-02	8.83E-02
ENSTTRP00000010668	<i>FAM131A</i>	3.35E-03	2.57E-02	ENSTTRP00000016245	<i>RBM41</i>	2.54E-02	8.92E-02
ENSTTRP00000014004	<i>CD226</i>	3.38E-03	2.58E-02	ENSTTRP00000011733	<i>PRDX4</i>	2.54E-02	8.92E-02
ENSTTRP00000006004	<i>POLR1C</i>	3.45E-03	2.63E-02	ENSTTRP00000016613	<i>PSMB11</i>	2.55E-02	8.92E-02
ENSTTRP00000010834	<i>TMEM82</i>	3.50E-03	2.65E-02	ENSTTRP00000008438	<i>FBXO24</i>	2.55E-02	8.92E-02
ENSTTRP00000004350	<i>DNAJC28</i>	3.54E-03	2.67E-02	ENSTTRP00000006030	<i>SLC13A1</i>	2.55E-02	8.92E-02
ENSTTRP00000008762	<i>STAB2</i>	3.58E-03	2.69E-02	ENSTTRP00000013784	<i>C1orf50</i>	2.56E-02	8.94E-02
ENSTTRP00000001914	<i>UIMC1</i>	3.60E-03	2.69E-02	ENSTTRP00000013446	<i>IL9</i>	2.58E-02	8.97E-02
ENSTTRP00000011833	<i>GSR</i>	3.64E-03	2.71E-02	ENSTTRP00000000166	<i>KIAA1407</i>	2.58E-02	8.97E-02
ENSTTRP00000002495	<i>CD48</i>	3.65E-03	2.71E-02	ENSTTRP00000007609	<i>NPM2</i>	2.60E-02	9.01E-02
ENSTTRP00000000278	<i>ASB7</i>	3.70E-03	2.74E-02	ENSTTRP00000004334	<i>SNRPA1</i>	2.60E-02	9.01E-02
ENSTTRP000000009456	<i>ARID5B</i>	3.73E-03	2.75E-02	ENSTTRP00000007689	<i>DDX59</i>	2.61E-02	9.02E-02
ENSTTRP00000009730	<i>NUDT22</i>	3.75E-03	2.75E-02	ENSTTRP00000005918	<i>KIF18B</i>	2.62E-02	9.04E-02
ENSTTRP00000012094	<i>SLC39A14</i>	3.77E-03	2.75E-02	ENSTTRP00000015077	<i>RRM2B</i>	2.63E-02	9.06E-02
ENSTTRP00000015092	<i>RHBDF2</i>	3.79E-03	2.76E-02	ENSTTRP00000001216	<i>GP1BA</i>	2.64E-02	9.08E-02
ENSTTRP00000010821	<i>NEIL1</i>	3.83E-03	2.77E-02	ENSTTRP00000001046	<i>GPATCH4</i>	2.66E-02	9.11E-02
ENSTTRP00000015799	<i>THOC1</i>	3.85E-03	2.77E-02	ENSTTRP00000014298	<i>CLSPN</i>	2.66E-02	9.11E-02
ENSTTRP00000003411	<i>CYP2R1</i>	3.88E-03	2.77E-02	ENSTTRP00000011415	<i>PARK2</i>	2.67E-02	9.11E-02
ENSTTRP00000008805	<i>C1orf222</i>	3.89E-03	2.77E-02	ENSTTRP00000013361	<i>GCFC2</i>	2.67E-02	9.11E-02
ENSTTRP00000008344	<i>N/A</i>	3.92E-03	2.77E-02	ENSTTRP00000016439	<i>DALRD3</i>	2.69E-02	9.14E-02
ENSTTRP00000002125	<i>HTATSF1</i>	3.94E-03	2.77E-02	ENSTTRP00000010520	<i>PPP1R32</i>	2.69E-02	9.14E-02
ENSTTRP00000010904	<i>INO80</i>	3.95E-03	2.77E-02	ENSTTRP00000014271	<i>COLAA4</i>	2.72E-02	9.23E-02
ENSTTRP00000013336	<i>B3V280_TURTR</i>	3.96E-03	2.77E-02	ENSTTRP00000008196	<i>WBSCR28</i>	2.77E-02	9.38E-02
ENSTTRP00000006594	<i>IL3</i>	3.96E-03	2.77E-02	ENSTTRP00000002809	<i>STARD10</i>	2.80E-02	9.47E-02
ENSTTRP00000011480	<i>MAPKAPK2</i>	3.97E-03	2.77E-02	ENSTTRP00000001050	<i>XCR1</i>	2.84E-02	9.58E-02
ENSTTRP00000011444	<i>SEC24C</i>	3.97E-03	2.77E-02	ENSTTRP00000016006	<i>DHX36</i>	2.89E-02	9.71E-02
ENSTTRP00000009311	<i>MDH1B</i>	3.99E-03	2.77E-02	ENSTTRP00000005027	<i>SPN</i>	2.90E-02	9.71E-02
ENSTTRP00000010004	<i>BRCA2</i>	4.04E-03	2.79E-02	ENSTTRP00000011529	<i>SETX</i>	2.90E-02	9.71E-02
ENSTTRP00000001395	<i>RINT1</i>	4.05E-03	2.79E-02	ENSTTRP00000007165	<i>C2orf49</i>	2.91E-02	9.71E-02
ENSTTRP00000015927	<i>C12orf66</i>	4.10E-03	2.81E-02	ENSTTRP00000000628	<i>OLFM4</i>	2.91E-02	9.71E-02
ENSTTRP00000002760	<i>LAMP2</i>	4.12E-03	2.82E-02	ENSTTRP00000011862	<i>CD6</i>	2.92E-02	9.71E-02
ENSTTRP00000005418	<i>DTNA</i>	4.14E-03	2.82E-02	ENSTTRP00000011554	<i>PRG4</i>	2.92E-02	9.71E-02
ENSTTRP00000010818	<i>ARHGAP35</i>	4.18E-03	2.84E-02	ENSTTRP00000015115	<i>CCDC122</i>	2.92E-02	9.71E-02
ENSTTRP00000015143	<i>CCDC141</i>	4.20E-03	2.84E-02	ENSTTRP00000004268	<i>BCAS3</i>	2.93E-02	9.72E-02
ENSTTRP00000012244	<i>SPATA19</i>	4.21E-03	2.84E-02	ENSTTRP00000005007	<i>ZNF521</i>	2.96E-02	9.81E-02
ENSTTRP00000000104	<i>STK31</i>	4.25E-03	2.85E-02	ENSTTRP00000010083	<i>BNIP1</i>	3.01E-02	9.92E-02
ENSTTRP00000015182	<i>GUF1</i>	4.36E-03	2.92E-02	ENSTTRP00000000526	<i>FAM43A</i>	3.01E-02	9.92E-02
ENSTTRP00000002595	<i>SNX16</i>	4.41E-03	2.93E-02	ENSTTRP00000005357	<i>SIGMARI</i>	3.01E-02	9.92E-02
ENSTTRP00000016090	<i>FOXP3</i>	4.42E-03	2.93E-02	ENSTTRP00000014455	<i>GCM1</i>	3.04E-02	1.00E-01
ENSTTRP00000016453	<i>CCDC36</i>	4.50E-03	2.98E-02	ENSTTRP00000010107	<i>THTPA</i>	3.05E-02	1.00E-01
ENSTTRP00000005020	<i>CXorf65</i>	4.52E-03	2.98E-02	ENSTTRP00000007774	<i>PGAP3</i>	3.06E-02	1.00E-01
ENSTTRP00000000977	<i>CRY2</i>	4.55E-03	2.99E-02	ENSTTRP00000001857	<i>N/A</i>	3.07E-02	1.00E-01
ENSTTRP00000007331	<i>JAK3</i>	4.67E-03	3.05E-02	ENSTTRP00000000180	<i>STT3A</i>	3.07E-02	1.00E-01
ENSTTRP00000015288	<i>FAM65B</i>	4.69E-03	3.05E-02				

**Supplementary Table 39 | Positively selected genes (PSGs) present in the minke whale and bottlenose dolphin.** A total of 64 PSGs were present in both whale genomes. N/A indicates not available gene name.

Protein ID (minke whale)	Protein ID (bottlenose dolphin)	Gene name (minke whale)	Gene name (bottlenose dolphin)	Minke whale		Bottlenose dolphin	
				P-value	Q-value	P-value	Q-value
BACU017034	ENSTTRP00000009852	<i>Col1a2</i>	<i>COL3A1</i>	<1.00E-11	<1.00E-08	2.88E-06	2.01E-04
BACU020345	ENSTTRP00000014813	<i>C4BPA</i>	<i>C4BPA</i>	7.13E-06	5.95E-04	2.39E-02	8.57E-02
BACU016007	ENSTTRP00000002191	<i>SPINK5</i>	<i>SPINK5</i>	1.55E-05	1.04E-03	1.10E-02	5.42E-02
BACU019534	ENSTTRP00000008565	<i>DPM1</i>	<i>DPM1</i>	2.02E-05	1.23E-03	5.90E-05	1.88E-03
BACU008820	ENSTTRP00000015646	<i>KIAA1671</i>	N/A	2.03E-05	1.23E-03	5.93E-03	3.58E-02
BACU018907	ENSTTRP00000004221	<i>C2orf72</i>	<i>C2orf73</i>	4.64E-05	2.29E-03	9.57E-03	4.96E-02
BACU000562	ENSTTRP00000011581	<i>Tdrd12</i>	<i>TD RD12</i>	6.97E-05	2.84E-03	9.02E-03	4.77E-02
BACU019689	ENSTTRP00000009730	<i>NUDT22</i>	<i>NUDT22</i>	7.50E-05	2.92E-03	3.75E-03	2.75E-02
BACU020364	ENSTTRP00000014985	<i>PRAP1</i>	<i>PRAP1</i>	1.02E-04	3.80E-03	1.98E-03	1.90E-02
BACU019919	ENSTTRP00000011586	<i>CEP89</i>	<i>CEP89</i>	1.05E-04	3.83E-03	8.80E-06	4.71E-04
BACU008502	ENSTTRP00000002390	<i>SPTBN5</i>	<i>SPTBN5</i>	1.10E-04	3.93E-03	4.69E-06	2.94E-04
BACU000213	ENSTTRP00000004453	<i>C1orf87</i>	<i>C1orf87</i>	1.22E-04	4.16E-03	1.03E-05	5.03E-04
BACU019914	ENSTTRP00000011565	<i>CCDC17</i>	<i>CCDC17</i>	1.26E-04	4.16E-03	1.25E-04	3.29E-03
BACU018780	ENSTTRP00000003398	<i>SLC6A3</i>	<i>SLC6A3</i>	1.98E-04	6.09E-03	9.54E-03	4.96E-02
BACU001324	ENSTTRP00000011444	<i>SEC24C</i>	<i>SEC24C</i>	3.08E-04	8.70E-03	3.97E-03	2.77E-02
BACU007890	ENSTTRP00000007588	<i>FAS</i>	<i>FAS</i>	3.55E-04	9.87E-03	3.82E-04	7.16E-03
BACU006841	ENSTTRP00000006030	<i>SLC13A1</i>	<i>SLC13A1</i>	4.77E-04	1.20E-02	2.55E-02	8.92E-02
BACU017483	ENSTTRP00000013449	<i>AKAP12</i>	<i>AKAP12</i>	4.81E-04	1.20E-02	7.28E-03	4.18E-02
BACU003063	ENSTTRP00000004537	<i>LPCAT3</i>	<i>LPCAT3</i>	5.11E-04	1.26E-02	1.66E-02	6.83E-02
BACU002213	ENSTTRP00000002248	<i>LRRK2</i>	<i>LRRK2</i>	5.44E-04	1.31E-02	1.66E-02	6.83E-02
BACU011924	ENSTTRP00000005027	<i>SPN</i>	<i>SPN</i>	5.96E-04	1.31E-02	2.90E-02	9.71E-02
BACU009330	ENSTTRP00000005513	<i>WDR52</i>	<i>WDR52</i>	6.01E-04	1.31E-02	1.18E-02	5.68E-02
BACU003792	ENSTTRP00000000526	<i>FAM43A</i>	<i>FAM43A</i>	6.50E-04	1.37E-02	3.01E-02	9.92E-02
BACU002570	ENSTTRP00000009850	<i>Khsrp</i>	<i>KHSRP</i>	6.74E-04	1.38E-02	4.48E-05	1.54E-03
BACU006439	ENSTTRP00000000206	<i>ANKRD40</i>	<i>ANKRD40</i>	7.79E-04	1.52E-02	1.52E-02	2.09E-02
BACU001239	ENSTTRP00000013530	<i>Polr2a</i>	<i>POLR2A</i>	8.88E-04	1.68E-02	2.94E-07	3.56E-05
BACU007621	ENSTTRP00000012859	<i>PPYR1</i>	<i>PPYR1</i>	8.93E-04	1.68E-02	2.18E-03	2.04E-02
BACU005661	ENSTTRP00000001154	<i>PDPN</i>	<i>PDPN</i>	9.47E-04	1.74E-02	1.23E-03	1.42E-02
BACU018235	ENSTTRP00000007863	<i>RABAC1</i>	<i>RABAC1</i>	1.20E-03	2.04E-02	1.03E-04	2.84E-03
BACU012013	ENSTTRP00000008196	<i>WBSCR28</i>	<i>WBSCR28</i>	1.35E-03	2.13E-02	2.77E-02	9.38E-02
BACU003049	ENSTTRP00000016063	<i>BRF2</i>	<i>BRF2</i>	1.47E-03	2.23E-02	1.88E-03	1.85E-02
BACU016680	ENSTTRP00000002342	<i>GLT25D1</i>	<i>FAM129C</i>	1.50E-03	2.23E-02	1.92E-04	4.36E-03
BACU013767	ENSTTRP00000009446	<i>NRP1</i>	<i>NRP1</i>	1.69E-03	2.45E-02	1.24E-02	5.81E-02
BACU009286	ENSTTRP000000010920	<i>GOLGB1</i>	<i>GOLGB1</i>	1.95E-03	2.73E-02	5.28E-04	8.89E-03
BACU001910	ENSTTRP00000009277	<i>P4HA1</i>	<i>P4HA1</i>	1.97E-03	2.74E-02	7.77E-04	1.09E-02
BACU019545	ENSTTRP00000008694	<i>DEC2R</i>	<i>DEC2R</i>	2.07E-03	2.78E-02	5.90E-04	9.33E-03
BACU019788	ENSTTRP00000010518	<i>PDE10A</i>	<i>PDE10A</i>	2.73E-03	3.47E-02	9.62E-04	1.23E-02
BACU004377	ENSTTRP00000000681	<i>RRP1B</i>	<i>RRP1B</i>	3.15E-03	3.83E-02	5.82E-03	3.55E-02
BACU011954	ENSTTRP00000014728	<i>PRSS53</i>	<i>PRSS53</i>	3.49E-03	4.13E-02	1.56E-02	6.60E-02
BACU009672	ENSTTRP000000011346	<i>SH3D21</i>	<i>SH3D21</i>	4.32E-03	4.72E-02	1.50E-02	6.43E-02
BACU011722	ENSTTRP00000004208	<i>RNASET2</i>	N/A	4.53E-03	4.75E-02	5.26E-03	3.33E-02
BACU003321	ENSTTRP00000009724	<i>SMTN</i>	<i>SMTN</i>	4.78E-03	4.90E-02	8.35E-03	4.56E-02
BACU003633	ENSTTRP00000010573	<i>CDHR5</i>	<i>CDHR5</i>	4.93E-03	4.96E-02	6.66E-07	7.33E-05
BACU014065	ENSTTRP00000015143	<i>CCDC141</i>	<i>CCDC141</i>	5.39E-03	5.28E-02	4.20E-03	2.84E-02
BACU018699	ENSTTRP00000002847	<i>AAGAB</i>	<i>AAGAB</i>	5.65E-03	5.35E-02	5.39E-03	3.37E-02
BACU008948	ENSTTRP00000000417	<i>AMBRA1</i>	<i>AMBRA1</i>	6.64E-03	5.93E-02	1.73E-02	6.98E-02
BACU000596	ENSTTRP00000014358	<i>DMKN</i>	<i>DMKN</i>	7.31E-03	6.25E-02	9.06E-06	4.71E-04
BACU020078	ENSTTRP00000012721	<i>DTX3L</i>	<i>DTX3L</i>	7.55E-03	6.39E-02	3.35E-03	2.57E-02
BACU011543	ENSTTRP00000015092	<i>RHBDF2</i>	<i>RHBDF2</i>	8.27E-03	6.90E-02	3.79E-03	2.76E-02
BACU013063	ENSTTRP00000006594	<i>IL3</i>	<i>IL3</i>	9.23E-03	7.35E-02	3.96E-03	2.77E-02
BACU008463	ENSTTRP00000009486	<i>CFP</i>	<i>CFP</i>	9.23E-03	7.35E-02	1.30E-02	6.00E-02
BACU009285	ENSTTRP00000010921	<i>Iqcb1</i>	<i>IQCB1</i>	9.30E-03	7.37E-02	1.74E-02	6.99E-02
BACU007776	ENSTTRP00000000577	<i>FMR1NB</i>	<i>FMR1NB</i>	9.73E-03	7.54E-02	2.38E-03	2.16E-02
BACU019926	ENSTTRP00000011634	<i>SEH1L</i>	<i>SEH1L</i>	1.01E-02	7.69E-02	6.90E-03	4.01E-02
BACU000447	ENSTTRP00000013856	<i>VEPH1</i>	<i>VEPH1</i>	1.06E-02	7.87E-02	4.84E-03	3.13E-02
BACU011539	ENSTTRP00000015087	<i>QRICH2</i>	<i>QRICH2</i>	1.11E-02	8.17E-02	1.02E-03	1.24E-02
BACU011209	ENSTTRP00000011733	<i>Prdx4</i>	<i>PRDX4</i>	1.17E-02	8.50E-02	2.54E-02	8.92E-02
BACU003261	ENSTTRP00000001740	<i>MED18</i>	<i>MED18</i>	1.31E-02	8.93E-02	1.47E-02	6.34E-02
BACU002504	ENSTTRP00000009041	<i>Icam4</i>	<i>ICAM4</i>	1.37E-02	9.13E-02	2.60E-04	5.44E-03
BACU003753	ENSTTRP00000000104	<i>STK31</i>	<i>STK31</i>	1.41E-02	9.25E-02	4.25E-03	2.85E-02
BACU018778	ENSTTRP00000003373	<i>SYCE2</i>	<i>SYCE2</i>	1.47E-02	9.57E-02	2.10E-06	1.62E-04
BACU004204	ENSTTRP00000013336	<i>IL2</i>	<i>B3V280_TURTR</i>	1.52E-02	9.79E-02	3.96E-03	2.77E-02
BACU004916	ENSTTRP00000005349	<i>SCML1</i>	<i>SCML1</i>	1.55E-02	9.87E-02	1.31E-03	1.49E-02
BACU013320	ENSTTRP00000012559	<i>AGA</i>	<i>AGA</i>	1.57E-02	9.93E-02	1.64E-02	6.79E-02

**Supplementary Table 40 | Pathway analysis of positively selected genes in the minke whale genome. P-values were calculated by Fisher's exact test.**

KEGG pathway	P-value	Q-value	Positively selected genes
Graft-versus-host disease	3.46E-04	3.00E-02	<i>FAS,IL1A,IL2</i>
Staphylococcus aureus infection	5.99E-04	3.00E-02	<i>DSG1,CFI,SELPLG</i>
Type I diabetes mellitus	1.45E-03	4.58E-02	<i>FAS,IL1A,IL2</i>
Glycerophospholipid metabolism	1.83E-03	4.58E-02	<i>CEPT1,DGKQ,LPCAT3,AGPAT1</i>
Glycerolipid metabolism	2.92E-03	5.84E-02	<i>DGAT1,DGKQ,AGPAT1</i>
Bladder cancer	3.62E-03	6.03E-02	<i>RASSF1,FIGF</i>
Pathogenic Escherichia coli infection	5.59E-03	7.99E-02	<i>LY96,ABLI</i>
Measles	6.49E-03	8.11E-02	<i>IFNAR2,RAB9A,FAS,IL1A,IL2</i>
Neurotrophin signaling pathway	1.08E-02	1.11E-01	<i>IRS4,SH2B3,ABLI,KIDINS220</i>
Salivary secretion	1.11E-02	1.11E-01	<i>ATP2B2,MUC7</i>
Fat digestion and absorption	1.48E-02	1.26E-01	<i>DGAT1,AGPAT1</i>
Complement and coagulation cascades	1.51E-02	1.26E-01	<i>KN1,CFI,C4BPA</i>
Glycosaminoglycan degradation	2.23E-02	1.70E-01	<i>HPSE</i>
Allograft rejection	2.38E-02	1.70E-01	<i>FAS,IL2</i>
Autoimmune thyroid disease	2.93E-02	1.95E-01	<i>FAS,IL2</i>
Retinol metabolism	3.24E-02	2.03E-01	<i>DGAT1</i>
Cytokine-cytokine receptor interaction	4.46E-02	2.07E-01	<i>IL3,IFNAR2,CCL22,FAS,FIGF,IL1A,IL2</i>
Apoptosis	4.72E-02	2.07E-01	<i>IL3,FAS,IL1A</i>
Protein digestion and absorption	4.91E-02	2.07E-01	<i>COL3A1,COL6A1</i>

**Supplementary Table 41 | Pathway analysis of positively selected genes in the bottlenose dolphin. P-values were calculated by Fisher's exact test.**

KEGG pathway	P-value	Q-value	Positively selected genes
Cytokine-cytokine receptor interaction	7.76E-05	1.22E-03	<i>IL3,IL18RAP,IL5,TNFRSF12A,PGF,IL9,CX3CL1,CXCL11,TNFRSF4,IL17RA,LEP,IL20RA,IL10RA,IL1RAP,CSF2RB,FAS,XCR1,GHR,IL2</i>
Complement and coagulation cascades	8.58E-04	6.73E-03	<i>PLAT,C9,CD59,C5,SERPINC1,C4BPA</i>
Jak-STAT signaling pathway	1.84E-03	8.04E-03	<i>LEP,IL3,EP300,IL5,IL20RA,IL10RA,IL9,SOS2,CSF2RB,JAK3,GHR,IL2</i>
Notch signaling pathway	2.05E-03	8.04E-03	<i>KAT2A,NOTCH2,EP300,DTX3L</i>
Asthma	2.71E-03	8.32E-03	<i>IL3,IL5,IL9</i>
Hematopoietic cell lineage	3.36E-03	8.32E-03	<i>CD38,IL3,CD37,IL5,CD44,CD59,GP1BA</i>
Protein digestion and absorption	3.71E-03	8.32E-03	<i>COL4A4,COL3A1,KCNK5,MEP1B,KCNJ13</i>
Fc epsilon RI signaling pathway	6.28E-03	1.23E-02	<i>IL3,IL5,SOS2,LCP2</i>
Circadian rhythm - mammal	7.23E-03	1.26E-02	<i>CRY2</i>
Lysosome	9.90E-03	1.31E-02	<i>AGA,LAMP2,CTSZ,NPC2,GUSB,GALC,AP4S1,CLN5</i>
Dorso-ventral axis formation	1.00E-02	1.31E-02	<i>NOTCH2,SOS2</i>
Drug metabolism - cytochrome P450	1.00E-02	1.31E-02	<i>FMO1,MAOA</i>
Natural killer cell mediated cytotoxicity	1.09E-02	1.32E-02	<i>CD48,ITGAL,SOS2,FAS,NFATC2,LCP2</i>
Allograft rejection	1.46E-02	1.36E-02	<i>IL5,FAS,IL2</i>
Pyrimidine metabolism	1.53E-02	1.36E-02	<i>DCTD,UPRT,AK3,POLR1C,RRM2B,POLR2B,CMPK2,POLR2A</i>
Staphylococcus aureus infection	1.65E-02	1.36E-02	<i>ITGAL,C5</i>
Graft-versus-host disease	1.65E-02	1.36E-02	<i>FAS,IL2</i>
Primary bile acid biosynthesis	1.65E-02	1.36E-02	<i>CYP46A1,AMACR</i>
Pentose and glucuronate interconversions	1.65E-02	1.36E-02	<i>GUSB,UGP2</i>
T cell receptor signaling pathway	1.83E-02	1.41E-02	<i>FOS,IL5,SOS2,NFATC2,IL2,LCP2</i>
Autoimmune thyroid disease	1.98E-02	1.41E-02	<i>IL5,FAS,IL2</i>
VEGF signaling pathway	1.98E-02	1.41E-02	<i>SH2D2A,MAPKAPK2,NFATC2</i>
Prion diseases	2.47E-02	1.69E-02	<i>C9,C5</i>
Starch and sucrose metabolism	3.48E-02	2.15E-02	<i>GUSB,UGP2</i>
Vitamin digestion and absorption	3.48E-02	2.15E-02	<i>SCARB1,LMBRD1</i>
Cell adhesion molecules (CAMs)	3.71E-02	2.15E-02	<i>ITGAL,CD274,CD6,CD226,SPN</i>
Thiamine metabolism	3.87E-02	2.15E-02	<i>THTPA</i>
Glycosphingolipid biosynthesis - ganglio series	3.87E-02	2.15E-02	<i>ST6GALNAC6</i>
Glutamatergic synapse	4.03E-02	2.15E-02	<i>GRM6,GRIK4,SLC38A1,GRIN3A</i>
Acute myeloid leukemia	4.11E-02	2.15E-02	<i>SOS2,PML,LEF1</i>
Systemic lupus erythematosus	4.67E-02	2.20E-02	<i>C9,C5</i>
Salivary secretion	4.67E-02	2.20E-02	<i>CD38,GUCY1A3</i>
Pyruvate metabolism	4.67E-02	2.20E-02	<i>ME3,GLO1</i>
Renal cell carcinoma	4.76E-02	2.20E-02	<i>EP300,PGF,SOS2,RAPGEF1</i>

**Supplementary Table 42 | Enrichment of the Gene Ontology (GO) categories of positively selected genes in the minke whale reference genome. Only GO categories with >5 PSGs are shown. P-values were calculated by Fisher's exact test.**

GO ID	# of PSGs	P-value	Q-value	GO domain	GO description
GO:0071310	13	2.61E-03	1.83E-03	biological_process	cellular response to organic substance
GO:0006886	13	2.98E-03	2.04E-03	biological_process	intracellular protein transport
GO:0016021	30	3.08E-03	2.11E-03	cellular_component	integral to membrane
GO:0022610	10	4.10E-03	2.72E-03	biological_process	biological adhesion
GO:0007155	10	4.10E-03	2.72E-03	biological_process	cell adhesion
GO:0015031	15	4.20E-03	2.78E-03	biological_process	protein transport
GO:0051716	53	4.69E-03	2.78E-03	biological_process	cellular response to stimulus
GO:0045184	15	5.22E-03	2.78E-03	biological_process	establishment of protein localization
GO:0070887	15	5.51E-03	2.78E-03	biological_process	cellular response to chemical stimulus
GO:0031224	30	5.58E-03	2.78E-03	cellular_component	intrinsic to membrane
GO:0008283	9	7.23E-03	3.43E-03	biological_process	cell proliferation
GO:0032101	8	7.44E-03	3.52E-03	biological_process	regulation of response to external stimulus
GO:0016337	6	8.32E-03	3.86E-03	biological_process	cell-cell adhesion
GO:0001819	6	8.32E-03	3.86E-03	biological_process	positive regulation of cytokine production
GO:0044425	47	8.47E-03	3.93E-03	cellular_component	membrane part
GO:0050896	63	8.59E-03	3.97E-03	biological_process	response to stimulus
GO:0046907	16	9.58E-03	4.18E-03	biological_process	intracellular transport
GO:0006163	8	1.18E-02	4.45E-03	biological_process	purine nucleotide metabolic process
GO:0001817	8	1.36E-02	5.02E-03	biological_process	regulation of cytokine production
GO:0042221	23	1.39E-02	5.09E-03	biological_process	response to chemical stimulus
GO:0051240	9	1.44E-02	5.23E-03	biological_process	positive regulation of multicellular organismal process
GO:0009150	7	1.44E-02	5.23E-03	biological_process	purine ribonucleotide metabolic process
GO:0019637	13	1.52E-02	5.51E-03	biological_process	organophosphate metabolic process
GO:0009259	7	1.80E-02	5.77E-03	biological_process	ribonucleotide metabolic process
GO:0007165	39	1.97E-02	6.22E-03	biological_process	signal transduction
GO:0072521	8	2.03E-02	6.39E-03	biological_process	purine-containing compound metabolic process
GO:0006955	6	2.03E-02	6.39E-03	biological_process	immune response
GO:0007166	21	2.15E-02	6.69E-03	biological_process	cell surface receptor signaling pathway
GO:0055085	11	2.20E-02	6.81E-03	biological_process	transmembrane transport
GO:0009117	9	2.58E-02	7.42E-03	biological_process	nucleotide metabolic process
GO:0009205	6	2.94E-02	8.22E-03	biological_process	purine ribonucleoside triphosphate metabolic process
GO:0051234	35	3.13E-02	8.62E-03	biological_process	establishment of localization
GO:0009199	6	3.15E-02	8.65E-03	biological_process	ribonucleoside triphosphate metabolic process
GO:0005730	12	3.18E-02	8.72E-03	cellular_component	nucleolus
GO:0051649	17	3.24E-02	8.86E-03	biological_process	establishment of localization in cell
GO:0006753	9	3.35E-02	8.94E-03	biological_process	nucleoside phosphate metabolic process
GO:0044212	6	3.35E-02	8.94E-03	molecular_function	transcription regulatory region DNA binding
GO:0006810	34	3.37E-02	8.94E-03	biological_process	transport
GO:1900542	6	3.37E-02	8.94E-03	biological_process	regulation of purine nucleotide metabolic process
GO:0009144	6	3.61E-02	9.11E-03	biological_process	purine nucleoside triphosphate metabolic process
GO:0006140	6	3.61E-02	9.11E-03	biological_process	regulation of nucleotide metabolic process
GO:0007167	7	3.67E-02	9.23E-03	biological_process	enzyme linked receptor protein signaling pathway
GO:0048646	10	3.87E-02	9.71E-03	biological_process	anatomical structure formation involved in morphogenesis
GO:0051049	12	4.02E-02	1.01E-02	biological_process	regulation of transport
GO:0001067	6	4.09E-02	1.02E-02	molecular_function	regulatory region nucleic acid binding
GO:0000975	6	4.09E-02	1.02E-02	molecular_function	regulatory region DNA binding
GO:0050789	93	4.20E-02	1.04E-02	biological_process	regulation of biological process
GO:0032879	17	4.60E-02	1.09E-02	biological_process	regulation of localization
GO:0009141	6	4.63E-02	1.10E-02	biological_process	nucleoside triphosphate metabolic process
GO:0055086	9	4.68E-02	1.10E-02	biological_process	nucleobase-containing small molecule metabolic process
GO:0008285	8	4.68E-02	1.10E-02	biological_process	negative regulation of cell proliferation
GO:0051046	6	4.92E-02	1.14E-02	biological_process	regulation of secretion

**Supplementary Table 43 | Enrichment of the Gene Ontology (GO) categories of positively selected genes in the bottlenose dolphin. Only GO categories with >14 PSGs are shown. P-values were calculated by Fisher's exact test.**

GO ID	# of PSGs	P-value	Q-value	GO domain	GO description
GO:0002682	28	2.37E-04	1.45E-04	biological_process	regulation of immune system process
GO:0001817	17	3.75E-04	2.28E-04	biological_process	regulation of cytokine production
GO:0050776	15	7.65E-04	4.21E-04	biological_process	regulation of immune response
GO:0033554	36	8.37E-04	4.58E-04	biological_process	cellular response to stress
GO:0043067	39	9.92E-04	5.35E-04	biological_process	regulation of programmed cell death
GO:0032101	16	1.07E-03	5.67E-04	biological_process	regulation of response to external stimulus
GO:0005615	20	1.30E-03	6.85E-04	cellular_component	extracellular space
GO:0010941	39	1.53E-03	8.00E-04	biological_process	regulation of cell death
GO:0042981	38	1.75E-03	8.68E-04	biological_process	regulation of apoptotic process
GO:0018193	18	1.96E-03	9.44E-04	biological_process	peptidyl-amino acid modification
GO:0022402	21	2.90E-03	1.35E-03	biological_process	cell cycle process
GO:0048584	32	3.17E-03	1.45E-03	biological_process	positive regulation of response to stimulus
GO:0006950	49	3.30E-03	1.49E-03	biological_process	response to stress
GO:0009790	15	4.34E-03	1.69E-03	biological_process	embryo development
GO:0048513	28	4.56E-03	1.69E-03	biological_process	organ development
GO:0016570	15	4.79E-03	1.69E-03	biological_process	histone modification
GO:0051716	96	5.02E-03	1.69E-03	biological_process	cellular response to stimulus
GO:0016569	15	5.28E-03	1.69E-03	biological_process	covalent chromatin modification
GO:0007165	73	5.65E-03	1.69E-03	biological_process	signal transduction
GO:0048583	59	6.19E-03	1.70E-03	biological_process	regulation of response to stimulus
GO:0002376	24	6.91E-03	1.90E-03	biological_process	immune system process
GO:0006974	23	6.92E-03	1.90E-03	biological_process	response to DNA damage stimulus
GO:0071840	83	7.16E-03	1.96E-03	biological_process	cellular component organization or biogenesis
GO:0022610	16	7.45E-03	2.02E-03	biological_process	biological adhesion
GO:0007155	16	7.45E-03	2.02E-03	biological_process	cell adhesion
GO:0043069	21	8.45E-03	2.18E-03	biological_process	negative regulation of programmed cell death
GO:0043066	21	8.45E-03	2.18E-03	biological_process	negative regulation of apoptotic process
GO:0016043	82	8.61E-03	2.21E-03	biological_process	cellular component organization
GO:0010033	31	8.67E-03	2.23E-03	biological_process	response to organic substance
GO:0051239	44	9.69E-03	2.47E-03	biological_process	regulation of multicellular organismal process
GO:0006464	60	1.14E-02	2.83E-03	biological_process	cellular protein modification process
GO:0036211	60	1.14E-02	2.83E-03	biological_process	protein modification process
GO:0060548	21	1.19E-02	2.94E-03	biological_process	negative regulation of cell death
GO:0043412	63	1.20E-02	2.95E-03	biological_process	macromolecule modification
GO:0016568	16	1.22E-02	2.99E-03	biological_process	chromatin modification
GO:0044421	20	1.24E-02	3.04E-03	cellular_component	extracellular region part
GO:0006325	18	1.26E-02	3.07E-03	biological_process	chromatin organization
GO:0071841	70	1.30E-02	3.11E-03	biological_process	cellular component organization or biogenesis at cellular level
GO:0050793	37	1.32E-02	3.14E-03	biological_process	regulation of developmental process
GO:0048856	53	1.48E-02	3.35E-03	biological_process	anatomical structure development
GO:0071842	69	1.56E-02	3.35E-03	biological_process	cellular component organization at cellular level
GO:0044699	50	1.60E-02	3.35E-03	biological_process	single-organism process
GO:0050896	116	1.75E-02	3.36E-03	biological_process	response to stimulus
GO:0009891	37	2.39E-02	4.32E-03	biological_process	positive regulation of biosynthetic process
GO:0051276	23	2.45E-02	4.39E-03	biological_process	chromosome organization
GO:0007166	36	2.50E-02	4.47E-03	biological_process	cell surface receptor signaling pathway
GO:0006281	15	2.69E-02	4.75E-03	biological_process	DNA repair
GO:0071310	19	2.72E-02	4.75E-03	biological_process	cellular response to organic substance
GO:0031328	36	2.82E-02	4.89E-03	biological_process	positive regulation of cellular biosynthetic process
GO:0045595	24	2.91E-02	4.97E-03	biological_process	regulation of cell differentiation
GO:2000026	27	2.99E-02	4.97E-03	biological_process	regulation of multicellular organismal development
GO:0035556	28	3.01E-02	4.97E-03	biological_process	intracellular signal transduction
GO:0010647	23	3.01E-02	4.97E-03	biological_process	positive regulation of cell communication
GO:0009967	22	3.28E-02	5.01E-03	biological_process	positive regulation of signal transduction
GO:0050794	174	3.36E-02	5.08E-03	biological_process	regulation of cellular process
GO:0038023	16	3.36E-02	5.08E-03	molecular_function	signaling receptor activity
GO:0045944	21	3.39E-02	5.12E-03	biological_process	positive regulation of transcription from RNA polymerase II promoter
GO:0006996	48	3.41E-02	5.13E-03	biological_process	organelle organization
GO:0048519	75	3.71E-02	5.45E-03	biological_process	negative regulation of biological process
GO:0045935	34	3.76E-02	5.52E-03	biological_process	positive regulation of nucleobase-containing compound metabolic process
GO:0051130	16	4.13E-02	5.95E-03	biological_process	positive regulation of cellular component organization
GO:0031325	49	4.16E-02	5.99E-03	biological_process	positive regulation of cellular metabolic process
GO:0004872	21	4.21E-02	6.06E-03	molecular_function	receptor activity
GO:0080134	19	4.24E-02	6.09E-03	biological_process	regulation of response to stress
GO:0051173	34	4.37E-02	6.27E-03	biological_process	positive regulation of nitrogen compound metabolic process
GO:0023056	22	4.40E-02	6.30E-03	biological_process	positive regulation of signaling
GO:0044707	39	4.85E-02	6.67E-03	biological_process	single-multicellular organism process
GO:0048522	78	4.95E-02	6.67E-03	biological_process	positive regulation of cellular process

**Supplementary Table 44 | Identification of rapidly evolving GO categories ( $P < 0.01$ ) in the minke whale.** A total of 25 GO categories with the P-values of less than 0.01 were determined.

GO ID	Minke whale $d_N/d_S$	Bottlenose dolphin $d_N/d_S$	GO description
GO:0006812	0.21991	0.159832	cation transport
GO:0016607	0.161337	0.127324	nuclear speck
GO:0005764	0.29403	0.229862	lysosome
GO:0042384	0.295328	0.294931	cilium assembly
GO:0005765	0.290197	0.2794	lysosomal membrane
GO:0016032	0.182753	0.157731	viral reproduction
GO:0035556	0.174035	0.169993	intracellular signal transduction
GO:0016023	0.217754	0.190485	cytoplasmic membrane-bounded vesicle
GO:0016772	0.229162	0.213285	transferase activity transferring phosphorus-containing groups
GO:0006457	0.235554	0.165971	protein folding
GO:0004672	0.215784	0.197361	protein kinase activity
GO:0055114	0.246857	0.219026	oxidation-reduction process
GO:0007267	0.243064	0.228601	cell-cell signaling
GO:0005739	0.263554	0.224436	mitochondrion
GO:0015758	0.220386	0.200679	glucose transport
GO:0003735	0.355027	0.217014	structural constituent of ribosome
GO:0043065	0.263382	0.25184	positive regulation of apoptotic process
GO:0006464	0.156526	0.146669	cellular protein modification process
GO:0042127	0.250692	0.19339	regulation of cell proliferation
GO:0008380	0.154994	0.13836	RNA splicing
GO:0004713	0.217155	0.184802	protein tyrosine kinase activity
GO:0016773	0.194224	0.170727	phosphotransferase activity alcohol group as acceptor
GO:0003899	0.153373	0.11257	DNA-directed RNA polymerase activity
GO:0055085	0.221358	0.205496	transmembrane transport
GO:0008237	0.275232	0.261561	metallopeptidase activity

**Supplementary Table 45 | Identification of slowly evolving GO categories ( $P < 0.01$ ) in the minke whale.** A total of 9 GO categories with the P-values of less than 0.01 were determined.

GO ID	Minke whale $d_N/d_S$	Bottlenose dolphin $d_N/d_S$	GO description
GO:0003779	0.051663	0.270064	actin binding
GO:0020037	0.163091	0.182875	heme binding
GO:0030054	0.041018	0.250722	cell junction
GO:0042802	0.039629	0.214656	identical protein binding
GO:0016311	0.00798	0.167447	dephosphorylation
GO:0007409	0.197477	0.20761	axonogenesis
GO:0016791	0.007955	0.16608	phosphatase activity
GO:0005737	0.058753	0.232748	cytoplasm
GO:0001726	0.017306	0.20791	ruffle

**Supplementary Table 46 | Copy number different genes in fin whale compared to minke whale genome.** The copy number differences in the fin whale were calculated by comparing its gene copy numbers to those in the minke whale genome.

Gene Name	Gene Full Name	Copy number
<i>AOX1</i>	aldehyde oxidase 1	Loss
<i>AP3S2</i>	adaptor-related protein complex 3, sigma 2 subunit	Loss
<i>ATXN1L</i>	ataxin 1-like	Loss
<i>AVPR1B</i>	arginine vasopressin receptor 1B	Loss
<i>AWAT2</i>	acyl-CoA wax alcohol acyltransferase 2	Loss
<i>C1orf186</i>	chromosome 1 open reading frame 186	Loss
<i>C1orf65</i>	chromosome 1 open reading frame 65	Loss
<i>C20orf72</i>	chromosome 20 open reading frame 72	Loss
<i>C2orf90</i>	chromosome 2 open reading frame 90	Loss
<i>CERBERUS</i>	cerberus 1, DAN family BMP antagonist	Loss
<i>Cbx1</i>	chromobox 1	Loss
<i>DUX1</i>	double homeobox 1	Gain
<i>EIF1AX</i>	eukaryotic translation initiation factor 1A, X-linked	Loss
<i>FAM170A</i>	family with sequence similarity 170, member A	Gain
<i>FAM3C</i>	family with sequence similarity 3, member C	Gain
<i>FAM72A</i>	family with sequence similarity 72, member A	Loss
<i>Glud1</i>	glutamate dehydrogenase 1	Loss
<i>H2BFWT</i>	H2B histone family, member W, testis-specific	Gain
<i>Hist2h2be</i>	histone cluster 2, H2be	Gain
<i>Hnrnpa3</i>	heterogeneous nuclear ribonucleoprotein A3	Gain
<i>IFI27</i>	interferon, alpha-inducible protein 27	Loss
<i>IFI27L2</i>	interferon, alpha-inducible protein 27-like 2	Loss
<i>IQCF3</i>	IQ motif containing F3	Loss
<i>KCNK13</i>	potassium channel, subfamily K, member 13	Loss
<i>KRCC1</i>	lysine-rich coiled-coil 1	Loss
<i>LRP12</i>	low density lipoprotein receptor-related protein 12	Gain
<i>Lin28a</i>	lin-28 homolog A (C. elegans)	Loss
<i>Mark2</i>	MAP/microtubule affinity-regulating kinase 2	Gain
<i>NAV1</i>	neuron navigator 1	Loss
<i>NIPA2</i>	non imprinted in Prader-Willi/Angelman syndrome 2	Loss
<i>Nutf2</i>	nuclear transport factor 2	Loss
<i>OR10C1</i>	olfactory receptor, family 10, subfamily C, member 1	Loss
<i>OR10H1</i>	olfactory receptor, family 10, subfamily H, member 1	Loss
<i>OR10H4</i>	olfactory receptor, family 10, subfamily H, member 4	Loss
<i>OR2A2</i>	olfactory receptor, family 2, subfamily A, member 2	Loss
<i>OR2B2</i>	olfactory receptor, family 2, subfamily B, member 2	Loss
<i>OR2G3</i>	olfactory receptor, family 2, subfamily G, member 3	Loss
<i>OR5V1</i>	olfactory receptor, family 5, subfamily V, member 1	Loss
<i>OR8A1</i>	olfactory receptor, family 8, subfamily A, member 1	Loss
<i>OR8S1</i>	olfactory receptor, family 8, subfamily S, member 1	Loss
<i>Olf481</i>	olfactory receptor 481	Loss
<i>PPP4R4</i>	protein phosphatase 4, regulatory subunit 4	Loss
<i>Phf5a</i>	PHD finger protein 5A	Loss
<i>Pol</i>	pol protein	Loss
<i>RFFL</i>	ring finger and FYVE-like domain containing E3 ubiquitin protein ligase	Loss
<i>S6PDH</i>	sorbitol 6-phosphate dehydrogenase	Loss
<i>SCAMP1</i>	secretory carrier membrane protein 1	Loss
<i>SCTR</i>	secretin receptor	Loss
<i>SEC31A</i>	SEC31 homolog A (S. cerevisiae)	Gain
<i>SERPINB13</i>	serpin peptidase inhibitor, clade B (ovalbumin), member 13	Loss
<i>SLC25A2</i>	solute carrier family 25 (mitochondrial carrier; ornithine transporter) member 2	Loss
<i>SPDYE4</i>	speedy/RINGO cell cycle regulator family member E4	Gain
<i>SPTLC1</i>	serine palmitoyltransferase, long chain base subunit 1	Loss
<i>SRY</i>	sex determining region Y	Loss
<i>SUSD4</i>	sushi domain containing 4	Loss
<i>SVIP</i>	small VCP/p97-interacting protein	Loss
<i>TADA2A</i>	transcriptional adaptor 2A	Loss
<i>TAF7</i>	TAF7 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 55kDa	Loss
<i>TIMM8A</i>	translocase of inner mitochondrial membrane 8 homolog A (yeast)	Gain
<i>TMEM37</i>	transmembrane protein 37	Loss
<i>TSPY1</i>	testis specific protein, Y-linked 1	Loss
<i>TSPY10</i>	testis specific protein, Y-linked 10	Loss
<i>TSPY3</i>	testis specific protein, Y-linked 3	Loss
<i>VPS29</i>	vacuolar protein sorting 29 homolog (S. cerevisiae)	Loss
<i>ZNF791</i>	zinc finger protein 791	Loss
<i>marinerT</i>	transposase mariner transposase undefined product (IC)	Gain
<i>ppe24</i>	PPE family protein PPE24	Gain

**Supplementary Table 47 | Copy number different genes in finless porpoise compared to bottlenose dolphin genome.** The copy number differences in the finless porpoise were calculated by comparing its gene copy numbers to those in the bottlenose dolphin genome.

Gene Name	Gene Full Name	Copy number
<i>5S_rRNA</i>	5S ribosomal RNA	Loss/Gain
<i>5_8S_rRNA</i>	5.8S ribosomal RNA	Loss/Gain
<i>AIBG</i>	alpha-1-B glycoprotein	Loss
<i>ACOT6</i>	acyl-CoA thioesterase 6	Loss
<i>CMA1</i>	chymase 1, mast cell	Loss
<i>FAM75D1</i>	family with sequence similarity 75, member D1	Loss
<i>FAM84B</i>	family with sequence similarity 84, member B	Loss
<i>GBP2</i>	guanylate binding protein 2, interferon-inducible	Loss
<i>HOXC9</i>	homeobox C9	Loss
<i>KLRD1</i>	killer cell lectin-like receptor subfamily D, member 1	Loss
<i>LIPF</i>	lipase, gastric	Loss
<i>MRPL17</i>	mitochondrial ribosomal protein L17	Loss
<i>PLEKHJ1</i>	pleckstrin homology domain containing, family J member 1	Loss
<i>PRR9</i>	proline rich 9	Loss
<i>RNF187</i>	ring finger protein 187	Loss
<i>SNORA16</i>	small nucleolar RNA SNORA16B/SNORA16A family	Loss
<i>SNORA17</i>	small nucleolar RNA, H/ACA box 17	Loss
<i>SNORA18</i>	small nucleolar RNA, H/ACA box 18	Loss
<i>SNORA40</i>	small nucleolar RNA, H/ACA box 40	Loss
<i>SNORA43</i>	small nucleolar RNA, H/ACA box 43	Loss
<i>SNORA72</i>	small nucleolar RNA, H/ACA box 72	Loss
<i>SNORD112</i>	small nucleolar RNA, C/D box 112	Loss
<i>SNORD115</i>	small nucleolar RNA, C/D box 115 cluster	Loss
<i>SNORD116</i>	small nucleolar RNA, C/D box 116 cluster	Loss
<i>TGM5</i>	transglutaminase 5	Loss
<i>THEG</i>	theg spermatid protein	Loss
<i>TMEM37</i>	transmembrane protein 37	Gain
<i>TSR2</i>	TSR2, 20S rRNA accumulation, homolog (S. cerevisiae)	Loss
<i>U2</i>	tegument protein UL23	Loss
<i>U4</i>	protein UL27	Loss
<i>U5</i>	putative viroporin	Loss
<i>U6</i>	ncRNA	Loss
<i>U6atac</i>	U6atac snRNA at 29B	Loss
<i>ZNF77</i>	zinc finger protein 77	Loss
<i>ZNF852</i>	zinc finger protein 852	Loss
<i>snoZ30</i>	ncRNA	Loss

**Supplementary Table 48 | Enrichment of the Gene Ontology (GO) categories for copy number different genes in fin whale and finless porpoise.** P-values were calculated by Fisher's exact test. N/A indicates not available gene name.

Species	CNV	P-value	GO Category	Genes
Fin whale	Loss	1.71E-11	signal transducer activity	<i>AVPR1B, OR10C1, OR10H1, OR10H4, OR2A2, OR2B2, OR2G3, OR5V1, OR8A1, OR8S1, Olf481, SCTR</i>
		1.74E-06	signal transduction	<i>AVPR1B, OR10C1, OR10H1, OR10H4, OR2A2, OR2B2, OR2G3, OR5V1, OR8A1, OR8S1, Olf481, SCTR</i>
	Gain	2.66E-02	chromosome organization	<i>TSPY1, TSPY10, TSPY3</i>
		5.49E-03	chromosome organization	<i>H2BFWT, Hist2h2be</i>
		1.00E-02	mitochondrion organization	<i>TIMM8A</i>
		1.04E-02	macromolecular complex assembly	<i>H2BFWT, Hist2h2be</i>
		1.15E-02	membrane organization	<i>TIMM8A</i>
		1.22E-02	DNA binding	<i>DUX1, H2BFWT, Hist2h2be</i>
		1.25E-02	cellular component assembly	<i>H2BFWT, Hist2h2be</i>
		4.18E-02	protein targeting	<i>TIMM8A</i>
4.60E-02	kinase activity	<i>Mark2</i>		
Finless porpoise	Loss	3.86E-04	methyltransferase activity	N/A
		3.77E-03	extracellular region	<i>A1BG, N/A</i>
		3.18E-02	intracellular	<i>CMA1, MRPL17, N/A, ZNF77, ZNF852</i>
	Gain	1.85E-02	nucleocytoplasmic transport	N/A
		4.80E-02	structural molecule activity	<i>TMEM37</i>

**Supplementary Table 49 | qPCR results for the *PRDX1* and *OGT* genes.** The expansions of the *PRDX1* and *OGT* gene families in the minke and fin whales were validated using qPCR. *ACTB1* gene is a single copy gene.

Species	Gene name / conc.	0.2ng	2ng	20ng	Efficiency	Relative copy number
Minke whale	<i>PRDX1</i>	30.070	26.160	22.590	0.850886	5.730
	<i>OGT</i>	28.320	24.120	20.590	0.81439	22.90
	<i>ACTB1</i>	30.600	26.540	22.720	0.8	
Fin whale	<i>PRDX1</i>	31.460	27.620	23.800	0.824301	4.110
	<i>OGT</i>	29.640	25.540	21.620	0.775728	18.64
	<i>ACTB1</i>	30.960	26.960	23.010	0.784729	

Note: Efficiency (E) =  $10^{(-1/\text{slop})} - 1$

Relative copy number  $\Delta\Delta\text{Ct} = (\text{Target Ct} - \text{HKG Ct}) - (\text{single copy gene Ct} - \text{HKG Ct})$

**Supplementary Table 50 | Primer pairs used for the qPCR.** The expansions of the *PRDX1* and *OGT* gene families in the minke and fin whales were validated using qPCR. *ACTB1* gene is a single copy gene.

Scaffold	Target position	target gene	Left primer	Right primer
scaffold170	2181603	<i>PRDX1</i>	GATGGGGCAGAGGAACTTAAGAA	GGGGTCTGATATCAAGGGAATGT
scaffold718	5228	<i>OGT</i>	TCAAGAAATGCTTTGAGGTCGAT	GCCCCATACTTTCTTTATTGGTG
scaffold4	17138524	<i>ACTB1</i>	CAGATCATGTTCGAGACCTCAA	GTACCCCTCGTAGATAGGCACC

**Supplementary Table 51 | Occurrence of blood clotting-related genes in artiodactyl species.** The lack of the *KLKB1* gene (Fletcher factor) in the minke whale was confirmed by aligning human *KLKB1* to scaffolds (only 5/14 were confirmed; >80% identity and >50% coverage) and mapping the RNA-seq data.

	minke whale	bottlenose dolphin	cow	pig
<i>F11</i> (PTA)	Yes	Yes	Yes	Yes
<i>F12</i> (Hegeman factor)	Yes	Yes	Yes	Yes
<i>HGFAC</i>	Yes	Yes	Yes	Yes
<i>HGF</i>	Yes	Yes	Yes	Yes
<i>PLAT</i> (tissue-type plasminogen activator)	Yes	Yes	Yes	Yes
<i>F5</i>	Yes	Yes	Yes	Yes
<i>F7</i>	Yes	Yes	Yes	Yes
<i>F8</i>	Yes	Yes	Yes	Yes
<i>F9</i>	Yes	Yes	Yes	Yes
<i>F10</i>	Yes	Yes	Yes	Yes
<i>F2</i> (prothrombin)	Yes	Yes	Yes	Yes
<i>FGG</i> (fibrinogen)	Yes	Yes	Yes	Yes
<i>FGA</i> (fibrinogen)	Yes	Yes	Yes	Yes
<i>FGB</i> (fibrinogen)	Yes	Yes	Yes	Yes
<i>KLKB1, KLK3</i> (Fletcher factor)	<b>No</b>	<b>No</b>	Yes	Yes

**Supplementary Table 52 | Top 50 genes enriched in each organs.** Highly expressed organ-specific genes were detected using transcriptome data from the brain, heart, kidney, liver, lung, muscle01, muscle02, and muscle03 tissue samples. FPKM is fragments per kilobase of transcript per million mapped reads.

Brain		Heart		Kidney		Liver	
gene	FPKM	gene	FPKM	gene	FPKM	gene	FPKM
<i>SNAP25</i>	1921.94	<i>MYL2</i>	41145.1	<i>GPX3</i>	7644.94	<i>HP</i>	12943.5
<i>S100B</i>	1604.3	<i>STARD3</i>	17657	<i>PRAP1</i>	2563.59	<i>APOC3</i>	11644.6
<i>ALDOC</i>	1317.81	<i>TNNT2</i>	10769.7	<i>PDZK1IP1</i>	1385.24	<i>RBP4</i>	9867.9
<i>CKB</i>	1189.8	<i>CCDC12</i>	6834.69	<i>ATP1B1</i>	1238.29	<i>SERPINA1</i>	9397.41
<i>YWHAH</i>	1189.57	<i>DNAAF3</i>	5983.91	<i>ATPIA1</i>	1213.98	<i>APOA1</i>	8954.46
<i>STMN2</i>	1014.39	<i>KCNQ2</i>	3229.25	<i>KLK1</i>	1202.83	<i>FTL</i>	8933.79
<i>CPE</i>	609.106	<i>MYBPC3</i>	2957.44	<i>IGFBP7</i>	1174.34	<i>APOE</i>	6832.98
<i>MBP</i>	588.12	<i>CSRP3</i>	2877.37	<i>UMOD</i>	1119.26	<i>FGB</i>	5576.41
<i>SPARCL1</i>	563.543	<i>ACTC1</i>	2152.65	<i>TMBIM6</i>	1063.63	<i>APOC2</i>	4601.15
<i>PEA15</i>	532.124	<i>TRIM54</i>	1409.37	<i>GABARAPLI</i>	940.547	<i>HPD</i>	4533.45
<i>RTN1</i>	519.291	<i>MYOZ2</i>	1392.66	<i>S100G</i>	888.062	<i>FGA</i>	4092.5
<i>PVALB</i>	510.026	<i>MYOM2</i>	1182.69	<i>NDRG1</i>	870.952	<i>APOA2</i>	3801.84
<i>CBLN3</i>	500.5	<i>FABP3</i>	1091.27	<i>SPP1</i>	711.183	<i>GATM</i>	3232.42
<i>NDRG4</i>	496.615	<i>FHL1</i>	866.342	<i>LGMN</i>	695.932	<i>ICK</i>	2883.45
<i>ATP6V1G2</i>	451.591	<i>ACO2</i>	828.856	<i>ALDH7A1</i>	627.366	<i>SERPING1</i>	2672.7
<i>CBLN1</i>	432.201	<i>BRK1</i>	793.812	<i>CDH16</i>	623.787	<i>HPX</i>	2558.24
<i>GDI1</i>	431.288	<i>LPL</i>	737.94	<i>MME</i>	512.365	<i>CEP112</i>	2506.17
<i>ENO2</i>	423.58	<i>HSPB7</i>	695.37	<i>ABAT</i>	504.921	<i>FGG</i>	2333.82
<i>GPM6A</i>	419.943	<i>MYL9</i>	623.492	<i>AMN</i>	486.206	<i>SAA2</i>	2266.22
<i>CHGB</i>	413.948	<i>SMYD2</i>	621.492	<i>ALDH6A1</i>	458.185	<i>DAB2</i>	2094.65
<i>SYTI</i>	407.67	<i>RABAC1</i>	587.145	<i>PDZK1</i>	451.62	<i>IGFBP1</i>	2063.82
<i>ABCF3</i>	402.081	<i>HADHA</i>	538.579	<i>PNP</i>	447.738	<i>VTN</i>	2013.52
<i>UCHL1</i>	389.004	<i>CASQ2</i>	523.539	<i>SLC5A2</i>	443.986	<i>HRSP12</i>	2012.05
<i>VSNL1</i>	388.955	<i>ATL3</i>	510.291	<i>SLC5A10</i>	442.954	<i>C4B</i>	1937.52
<i>SNCA</i>	365.162	<i>LDHB</i>	445.128	<i>ABP1</i>	434.121	<i>BHMT</i>	1919.99
<i>TSPAN7</i>	351.915	<i>TESC</i>	410.015	<i>HAO2</i>	426.786	<i>ORM1</i>	1814.42
<i>GFAP</i>	351.204	<i>MYOM1</i>	391.724	<i>TXNDC17</i>	380.277	<i>ADH1B</i>	1808.66
<i>DKK3</i>	334.806	<i>ANKRD2</i>	381.095	<i>C9ORF71</i>	377.553	<i>ASS1</i>	1800.58
<i>HPCAL1</i>	315.51	<i>TMEM233</i>	334.367	<i>FAM151A</i>	360.739	<i>ATG12</i>	1692.34
<i>PCP4</i>	309.792	<i>AIFM1</i>	328.71	<i>AQP7</i>	360.063	<i>AMBP</i>	1603.25
<i>TMEM59L</i>	299.75	<i>HADHB</i>	307.14	<i>MAT2A</i>	336.069	<i>KNG2</i>	1382.98
<i>ZIC1</i>	296.61	<i>CDH13</i>	303.067	<i>UPB1</i>	331.262	<i>TF</i>	1350.58
<i>CIORF61</i>	294.41	<i>NHP2L1</i>	299.184	<i>SLC22A1</i>	318.083	<i>GC</i>	1347.19
<i>PPP2R1A</i>	291.161	<i>FHL2</i>	297.799	<i>ANXA8</i>	316.493	<i>FGL1</i>	1342.51
<i>PER1</i>	287.517	<i>CS</i>	294.188	<i>SLC1A5</i>	283.557	<i>MGST1</i>	1179.77
<i>ATP1B2</i>	283.071	<i>NPPB</i>	293.65	<i>MIOX</i>	279.465	<i>CYP3A28</i>	1143.82
<i>SLC1A3</i>	282.761	<i>HSPB6</i>	290.61	<i>CLDN2</i>	273.691	<i>RBP2</i>	1067.19
<i>KIF1B</i>	280.393	<i>PERP</i>	286.568	<i>PTGDS</i>	268.966	<i>ITIH1</i>	1017.06
<i>HTRA1</i>	279.914	<i>ERCC2</i>	271.516	<i>SLC34A1</i>	263.345	<i>CYP2E1</i>	1009.99
<i>CKMT1</i>	279.898	<i>KCNIP2</i>	245.283	<i>SLC37A4</i>	256.804	<i>GAMT</i>	1007.88
<i>APOPT1</i>	277.017	<i>SORBS2</i>	230.583	<i>SLC22A2</i>	256.645	<i>CFB</i>	1004.46
<i>ELMO1</i>	266.956	<i>MYZAP</i>	229.522	<i>EIF6</i>	255.172	<i>RC3H1</i>	1000.28
<i>ATPIA3</i>	263.116	<i>PKP2</i>	225.866	<i>EGF</i>	253.739	<i>HSD17B13</i>	995.509
<i>BCDIN3D</i>	262.51	<i>HSPB3</i>	214.531	<i>ACSM1</i>	250.134	<i>IGFBP2</i>	967.152
<i>STXBP1</i>	258.612	<i>TWF2</i>	198.753	<i>PEPD</i>	248.96	<i>HSD11B1</i>	913.783
<i>PLP1</i>	255.699	<i>NES</i>	196.118	<i>CTH</i>	241.587	<i>CPS1</i>	898.412
<i>DPYSL2</i>	253.59	<i>EGLN3</i>	195.812	<i>CUBN</i>	234.738	<i>TDO2</i>	883.913
<i>SPOCK2</i>	253.394	<i>CARHSP1</i>	191.062	<i>SLC36A2</i>	232.825	<i>A1BG</i>	871.645
<i>DNMT1</i>	252.832	<i>PDHA1</i>	189.674	<i>CA12</i>	224.892	<i>MAT1A</i>	808.395
<i>CADM3</i>	252.396	<i>PPIF</i>	180.576	<i>SLC13A3</i>	223.32	<i>GUCA2A</i>	793.614

Lung		Muscle01		Muscle02		Muscle03	
gene	FPKM	gene	FPKM	gene	FPKM	gene	FPKM
<i>TMSB10</i>	81318.3	<i>TRIM63</i>	2115.5	<i>PPP1R12C</i>	8365.38	<i>ALDOA</i>	60859.5
<i>HSPA1L</i>	5496.53	<i>MYOT</i>	1591.09	<i>RPS29</i>	4717.53	<i>HBA</i>	4953.5
<i>HSPA5</i>	2349.11	<i>FLNC</i>	1322.04	<i>COX7B</i>	4686.28	<i>NLRC3</i>	3456.51
<i>CD9</i>	2238.14	<i>HSPB8</i>	1040.18	<i>MYH7</i>	4392.08	<i>CA3</i>	2840.4
<i>HSP90AA1</i>	1848.59	<i>XIRP1</i>	907.652	<i>RPS24</i>	3943.45	<i>LDHA</i>	301.397
<i>HLA-A</i>	1482.83	<i>MYF6</i>	839.425	<i>COX17</i>	3723.18	<i>TSG101</i>	203.466
<i>TMSB4X</i>	1424.68	<i>ASB15</i>	792.564	<i>TTYH2</i>	3434.63	<i>COL1A1</i>	177.748
<i>SFTPB</i>	1341.58	<i>ZFAND5</i>	669.277	<i>ATP5E</i>	2684.15	<i>COL1A2</i>	154.007
<i>AQP5</i>	1325.45	<i>ABRA</i>	555.889	<i>TPM3</i>	2606.12	<i>C6ORF164</i>	146.156
<i>ACTG1</i>	1234.26	<i>ASB5</i>	457.411	<i>TNNI1</i>	2579.39	<i>MPZ</i>	81.3598
<i>CD74</i>	1218.08	<i>PNRC1</i>	364.272	<i>RPL22</i>	2268.98	<i>THBS4</i>	71.4436
<i>CIQB</i>	1207.41	<i>CSORF65</i>	273.893	<i>RPL36</i>	2192.44	<i>FMOD</i>	61.1958
<i>SCGB3A2</i>	1199.24	<i>ABCF2</i>	250.147	<i>NDRG2</i>	1833.23	<i>KY</i>	32.0254
<i>SPARC</i>	1034.29	<i>KBTBD5</i>	245.856	<i>TPM2</i>	1825.93	<i>SFRP4</i>	28.1626
<i>CIQA</i>	995.866	<i>USP28</i>	229.97	<i>RPS5</i>	1677.48	<i>TNMD</i>	24.1583
<i>TGAS113E22.1</i>	977.22	<i>GSPT1</i>	217.688	<i>MYBPH</i>	1655.41	<i>PMP2</i>	20.7635
<i>DNAJB1</i>	870.093	<i>VGLL2</i>	193.401	<i>MIDN</i>	1519.34	<i>SFRP2</i>	16.2094
<i>VIM</i>	842.803	<i>GPCPD1</i>	189.443	<i>RPL3L</i>	1474.66	<i>SLC25A33</i>	16.2079
<i>KRT18</i>	813.139	<i>ZNF622</i>	183.989	<i>DDIT4L</i>	1430.53	<i>MEOX2</i>	15.9301
<i>TAGLN</i>	733.477	<i>ALPK3</i>	182.058	<i>RPL18A</i>	1360.63	<i>GRHPR</i>	11.6722
<i>MYL6B</i>	729.158	<i>WWP1</i>	162.009	<i>NDUFA2</i>	1315.56	<i>LRRC30</i>	10.9911
<i>HSP90B1</i>	723.75	<i>FBXO40</i>	160.492	<i>MFSD7</i>	1206.79		
<i>TAGLN2</i>	722.325	<i>KAT2B</i>	155.303	<i>TUBA4A</i>	1194.58		
<i>PDPN</i>	718.474	<i>NFIL3</i>	154.762	<i>GYPC</i>	987.559		
<i>CIQC</i>	707.9	<i>PDE4D</i>	142.923	<i>NDUFA1</i>	909.681		
<i>TIMP3</i>	703.662	<i>MYH2</i>	133.915	<i>RPS20</i>	890.167		
<i>CCL23</i>	703.329	<i>HIST2H2BE</i>	133.08	<i>NDUFA3</i>	872.548		
<i>TPM1</i>	667.148	<i>GTF2B</i>	122.864	<i>TCEB2</i>	769.732		
<i>WFDC2</i>	649.585	<i>FBXO32</i>	122.829	<i>RPS19</i>	741.236		
<i>AGER</i>	642.788	<i>FXR1</i>	120.899	<i>PNPLA2</i>	737.281		
<i>S100A6</i>	635.137	<i>ARIH2</i>	118.638	<i>NDUFB1</i>	707.494		
<i>RHOA</i>	625.975	<i>PHKA1</i>	117.266	<i>PPDPF</i>	685.602		
<i>BMP1</i>	596.52	<i>OTUD1</i>	114.049	<i>GABARAP</i>	679.47		
<i>ICAM1</i>	569.573	<i>APBB2</i>	109.907	<i>EIF4EBP1</i>	655.614		
<i>CALR</i>	565.643	<i>RBFOX2</i>	106.87	<i>CACNG1</i>	609.557		
<i>MARCO</i>	563.729	<i>SYNPO2</i>	105.646	<i>FHL3</i>	593.462		
<i>CSRP1</i>	557.4	<i>KPNA3</i>	105.009	<i>MAP7D3</i>	591.496		
<i>CCL2</i>	556.335	<i>ARID5B</i>	102.889	<i>VAMP5</i>	585.551		
<i>SUPV3L1</i>	549.955	<i>NAA50</i>	95.6912	<i>IQUB</i>	583.917		
<i>ANXA2</i>	533.433	<i>FBXW7</i>	93.0084	<i>FAM162A</i>	581.817		
<i>IGH-VJ558</i>	531.954	<i>CRKL</i>	87.914	<i>AKIRIN1</i>	578.548		
<i>LYVE1</i>	527.595	<i>KPNA1</i>	87.712	<i>SHFM1</i>	540.869		
<i>IFI27</i>	525.785	<i>KLHL30</i>	86.0347	<i>RPL27A</i>	540.499		
<i>MGP</i>	520.076	<i>RBM20</i>	85.1711	<i>NDUFC2</i>	539.757		
<i>MCM5</i>	518.377	<i>PRKG1</i>	84.4479	<i>SGCA</i>	521.061		
<i>CD68</i>	507.606	<i>PPM1B</i>	79.326	<i>SNRNP25</i>	519.74		
<i>TIMP1</i>	503.92	<i>FOXO1</i>	74.0503	<i>RTN2</i>	517.695		
<i>SERPINH1</i>	451.931	<i>CLIP1</i>	73.5476	<i>DHRS7C</i>	516.779		
<i>RHOB</i>	448.527	<i>KLHL38</i>	72.6628	<i>RPL23A</i>	515.886		
<i>EPAS1</i>	441.347	<i>USP25</i>	70.8023	<i>RPL15</i>	475.126		

**Supplementary Table 53 | Top 10 pathway analysis results for enriched genes. P-values were calculated by Fisher's exact test.**

Pathway	# of genes	P-value	Pathway	# of genes	P-value
<b>Brain</b>			<b>Heart</b>		
Long-term potentiation	14	<0.000001	Dilated cardiomyopathy	9	<0.000001
Phosphatidylinositol signaling system	11	<0.000001	Vascular smooth muscle contraction	9	<0.000001
MAPK signaling pathway	38	<0.000001	Hypertrophic cardiomyopathy (HCM)	8	0.000001
Glutamatergic synapse	32	<0.000001	Cardiac muscle contraction	7	0.000002
Long-term depression	12	0.000002	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	6	0.000041
ErbB signaling pathway	13	0.000037	Nicotinate and nicotinamide metabolism	3	0.000045
Endocrine and other factor-regulated calcium reabsorption	8	0.000078	Gastric acid secretion	5	0.000049
Epithelial cell signaling in Helicobacter pylori infection	11	0.000128	Fatty acid elongation in mitochondria	2	0.000092
GnRH signaling pathway	12	0.000191	Citrate cycle (TCA cycle)	3	0.000267
Fc gamma R-mediated phagocytosis	11	0.000227	Pancreatic secretion	5	0.0003
<b>Kidney</b>			<b>Liver</b>		
D-Arginine and D-ornithine metabolism	1	<0.000001	Staphylococcus aureus infection	11	<0.000001
Lysine biosynthesis	3	<0.000001	Bile secretion	14	<0.000001
Mineral absorption	9	0.000001	Fat digestion and absorption	9	<0.000001
Protein digestion and absorption	11	0.000001	Complement and coagulation cascades	33	<0.000001
Amino sugar and nucleotide sugar metabolism	8	0.000008	PPAR signaling pathway	13	<0.000001
Vitamin digestion and absorption	6	0.000037	Primary bile acid biosynthesis	7	0.000001
Proximal tubule bicarbonate reclamation	4	0.000148	Histidine metabolism	6	0.000003
Ether lipid metabolism	4	0.000148	Arginine and proline metabolism	9	0.000005
Arginine and proline metabolism	7	0.000268	Drug metabolism - cytochrome P450	6	0.000008
Aldosterone-regulated sodium reabsorption	5	0.000453	Steroid hormone biosynthesis	8	0.000009
<b>Lung</b>			<b>Muscle01</b>		
Tuberculosis	25	<0.000001	Pathways in cancer	8	0.000033
Toxoplasmosis	21	<0.000001	Small cell lung cancer	4	0.000079
Malaria	13	<0.000001	Insulin signaling pathway	4	0.000204
Chagas disease (American trypanosomiasis)	23	<0.000001	MAPK signaling pathway	6	0.000243
Leukocyte transendothelial migration	19	<0.000001	Acute myeloid leukemia	3	0.000294
Rheumatoid arthritis	19	<0.000001	TGF-beta signaling pathway	3	0.000473
Osteoclast differentiation	30	<0.000001	Chronic myeloid leukemia	3	0.000547
Pathways in cancer	39	<0.000001	ErbB signaling pathway	3	0.000871
Leishmaniasis	14	0.000002	Thyroid cancer	2	0.000989
Toll-like receptor signaling pathway	17	0.000003	Viral myocarditis	2	0.002072
<b>Muscle02</b>			<b>Muscle03</b>		
Oxidative phosphorylation	13	<0.000001	Pyruvate metabolism	3	<0.000001
Cardiac muscle contraction	9	<0.000001	Glyoxylate and dicarboxylate metabolism	2	0.000007
Huntington's disease	13	<0.000001	Glycolysis / Gluconeogenesis	2	0.000085
Parkinson's disease	12	<0.000001	Nitrogen metabolism	1	0.000213
Alzheimer's disease	12	<0.000001	ECM-receptor interaction	2	0.000322
Ribosome	19	<0.000001	Pentose phosphate pathway	1	0.000795
Hypertrophic cardiomyopathy (HCM)	6	0.000236	Fructose and mannose metabolism	1	0.000997
Dilated cardiomyopathy	6	0.000372	Propanoate metabolism	1	0.001341
Fatty acid biosynthesis	1	0.002468	Wnt signaling pathway	2	0.001432
Valine, leucine and isoleucine biosynthesis	1	0.003663	Cysteine and methionine metabolism	1	0.001467

**Supplementary Table 54 | Neuroglobin and cytoglobin expression levels.** The expression levels of neuroglobin and cytoglobin were analyzed using RNA-seq data. The values are fragments per kilobase of transcript per million mapped reads (FPKM).

Tissue	Neuroglobin	Cytoglobin
Lung	0.0	12.65
Brain	2.64	39.40
Heart	0.12	3.77
Kidney	0.33	8.63
Liver	0.04	4.50
Muscle01	0	2.54
Muscle02	0	2.09
Muscle03	0	5.09

**Supplementary Table 55 | Read mapping statistics.** The filtered reads of six cetaceans were mapped to the minke whale genome.

Sample	# of All Pairs	# of Mapped Reads	# of Unmapped Reads	# of Mapped reads after removal of PCR duplicates	% of Mapped Reads	Mapping Depth
Minke whale01	1,032,532,282	961,846,499	70,685,783	772,080,137	93.15	33
Minke whale02	992,102,360	924,434,495	67,667,865	716,048,408	93.18	31
Minke whale03	827,706,704	770,733,110	56,973,594	686,925,436	93.12	30
Fin whale	1,157,484,176	952,460,024	205,024,152	706,301,952	82.29	30
Bottlenose dolphin	1,160,012,272	1,083,252,872	76,759,400	988,740,745	93.38	43
Finless porpoise	1,047,224,476	788,932,847	258,291,629	646,945,328	75.34	28

**Supplementary Table 56 | SNV and small indel statistics.** SNVs and indels of six cetaceans were found by mapping reads to the minke whale genome.

Sample	# of all Variant locus	# of homozygous SNV locus	# of heterozygous SNV locus	# of small indel locus
Minke whale01	3,036,996	1,272,900	1,369,779	394,317
Minke whale02	3,305,595	1,328,410	1,542,964	434,221
Minke whale03	3,389,176	1,372,913	1,587,639	428,624
Fin whale	34,195,413	27,679,344	3,693,762	2,822,307
Bottlenose dolphin	8,309,630	3,211,890	3,586,803	1,510,937
Finless porpoise	46,436,537	39,960,183	2,162,792	4,313,562

**Supplementary Table 57 | Estimation of the split times and mutation rates.** Dolphin, Whale, Cow, Pig, Dog, Human, Opossum, and Platypus are *Tursiops truncatus*, *Balaenoptera acutorostrata*, *Bos taurus*, *Sus scrofa*, *Canis lupus familiaris*, *Homo sapiens*, *Monodelphis domestica*, and *Ornithorhynchus anatinus*, respectively. CI means 95% credibility interval. The mutation rates were calculated by dividing the evolutionary distances in ML tree with the estimated split times. Since both of the ML tree and split times were computed on four-fold degenerate sites, the mutation rates represent the neutral substitution rate per site per year.

Split	Split Time (Mya)	CI for Split Time (Mya)	Branch length (ML Tree Distance)	Branch length (Million year, Split Time Tree)	Mutation rate (substitution per site per year)
Dolphin	-	-	0.025965	21.4571	1.2101E-09
Whale	-	-	0.022977	21.4571	1.0709E-09
Cow	-	-	0.135262	36.2121	3.7353E-09
Pig	-	-	0.132160	51.8773	2.5475E-09
Dog	-	-	0.150719	67.5253	2.2320E-09
Human	-	-	0.134345	96.7405	1.3887E-09
Opossum	-	-	0.440512	134.5189	3.2747E-09
Platypus	-	-	-	178.0198	-
Dolphin Whale	21.46	11.2-29.5	0.040650	14.7549	2.7550E-09
Dolphin Cow	36.21	28.8-43.9	0.018792	15.6652	1.1996E-09
Dolphin Pig	51.88	48.3-54.1	0.048403	15.648	3.0932E-09
Dolphin Dog	67.53	58.7-81.0	0.027581	29.2152	9.4406E-10
Dolphin Human	96.74	93.8-103.0	0.313184	37.7784	8.2900E-09
Dolphin Opossum	134.52	126.7-138.4	-	43.5009	-
Dolphin Platypus	178.02	162.3-191.2	-	-	-

## Supplementary Note

### Sequence filtering criteria

The following filtering criteria were applied to reduce the effects of sequencing errors on the assembly, thereby ensuring high quality reads.

- 1) Reads with ambiguous bases (represented by the letter N) or poly-A structures.
- 2) Reads with  $\geq 40\%$  low-quality bases (base quality  $\leq 7$ ) in small insert size libraries (170, 500, and 800 bp) and reads with  $\geq 60\%$  low-quality bases in large insert size libraries.
- 3) Reads with adapter contamination: Reads with  $\geq 10$  bp aligned to the adapter sequence ( $\leq 3$  bp mismatch allowed) were filtered out.
- 4) Small insert size reads in which read1 and read2 overlapped by  $\geq 10$  bp (10% mismatch allowed).
- 5) PCR duplications (reads were considered duplicates when read1 and read2 of the two paired-end reads were identical).

The genomes for comparative analyses (from three other minke whales, a fin whale, a bottlenose dolphin, and a finless porpoise) were filtered out as follows:

- 1) Reads with an N ratio  $> 10\%$ .
- 2) Reads with an average quality  $< 15$ .
- 3) Reads in which nucleotides with a quality of  $< 15$  comprise  $> 5\%$  of the read.
- 4) Reads containing at least  $> 15$ bp of an adapter sequence, with three mismatches allowed.
- 5) Reads containing PhiX sequences.

## Expansion of *PRDX1* and *OGT* homolog genes

### Quantitative PCR assay and amplification efficiency

Quantitative real-time PCR was performed using a Light Cycler 480 (Roche) with 1  $\mu$ l of gDNA template and 5  $\mu$ l of FastStart Universal SYBR Green Master Mix per reaction. The reaction volume was adjusted to 10  $\mu$ l with water and primer solutions.

The quantitative real-time PCR amplification program consisted of a 10-min pre-incubation at 95°C followed by 40 cycles of 95°C for 15 s and 55°C for 1 min. Independent experiments were performed in triplicate.

A melting curve was produced by increasing the temperature from 65°C to 97°C. The genomic DNA samples were diluted to produce 10-fold serial dilutions (0.2 ng, 2 ng, and 20 ng) to determine the amplification efficiency for the target genes (*PRDX1* and *OGT*) and HKG (*ACTB1*). The amplification efficiency of the respective genes was calculated as:

$$E=10^{(-1/\text{slope})} - 1.$$

### Determination of relative gene copy number

To calculate the copy number of the *PRDX1* and *OGT* genes in minke and fin whales, relative quantification of the samples was performed using the  $2^{-\Delta\Delta C_t}$  method<sup>1</sup>. The *FANCF* gene, which is known to be present in a single copy in whales, was used as a calibrator in this study. The *ACTB1* gene was also used as a housekeeping gene for expression normalization in all experiments. The calculation of  $\Delta\Delta C_t$  for the relative quantification of target genes (*PRDX1* and *OGT*) was performed as follows:

$$\Delta\Delta C_t = (\text{Target } C_t - \text{HKG } C_t) - (\text{single copy gene } C_t - \text{HKG } C_t).$$

## Olfactory receptor (OR) gene identification

We identified OR genes (functional, pseudogenes, or partial genes) in the minke whale genome. The number of functional OR genes and pseudogenes was much lower in the minke whale genome (n=134) compared with the genomes of other mammals, whereas the percentage of OR pseudogenes (n=74) was higher than that in other mammals<sup>2,3</sup> (~55%). The number of partial genes (n=117) with  $\geq 2$  OR motifs plus the number of pseudogenes (n=74) was much larger than the number of functional OR genes (n=60), which suggests that selection pressure was lost to maintain functional OR genes in the whale, most likely due to a lower reliance on olfaction for survival in a marine environment. This is consistent with the smaller size of the olfactory subgenomes and higher frequency of OR pseudogenization in underwater animals such as fish and frog.

tBLASTn was used to identify regions containing OR-related sequences with at least one of the following conserved motifs: MAYDRYVAIC (TMIII), KAFSTCASH (TMVI), PMLNPFYIY (TMVII), or variants thereof with <40% sequence variation from the conserved motifs. Sequences in the regions located 1 Kb upstream and downstream of the BLAST matches were selected from the regions identified. OR candidate sequences measuring 2 Kb in length were translated into amino acid sequences in all six frames. BLASTP was performed using 24,809 OR protein sequences from 222 species in NCBI and the translated OR candidate sequences to determine the positions of the start and stop codons of their open reading frames (ORFs) based on structural similarity to known OR proteins. For sequences that deviated from the sequences of reported OR proteins, the methionine and stop codons that were most similar in sequence context to those delineating the coding sequences of known OR proteins were selected as the start and end of the coding regions. The candidate sequences were considered “functional ORs” if they contained  $\geq 300$  amino acids without any interrupting stop codons and/or frameshifts in the ORFs, “OR pseudogenes” if they contained  $\geq 300$  amino acids with stop codons or frameshifts in the ORFs, and “partial ORs” if they <300 amino acids but matched the sequences of known OR genes. Sequences similar to non-OR G-protein-coupled receptors or partial sequences were removed from the analyses.

## Divergence time

Single copy gene families were used to construct a phylogenetic tree for *B. acutorostrata* and the other sequenced mammalian genomes. Four-fold degenerate sites were extracted from each family and concatenated to form one supergene for each species. The substitution model (HKY85+gamma) was selected, and PhyML v3.0<sup>4</sup> was used to reconstruct the phylogenetic tree.

The molecular clock for four-fold degenerate sites (neutral substitution rate per year) was estimated based on single copy gene families. The divergence time and neutral substitution rate per year (evolutionary distance/divergence time) among species were estimated.

The divergence time was estimated for the eight species using single copy gene families. Molecular sequence data of four-fold degenerate sites were used to estimate species divergence time using the program MCMCTREE v3.0<sup>5</sup> with an approximate likelihood calculation algorithm, which was implemented using the PAML package<sup>6</sup> (version 4.5). The fossil calibration times for the *Sus scrofa*-*Bos taurus* divergence (48.3–53.5 million years ago [MYA]), *Bos taurus*-*Homo sapiens* divergence (95.3–113 MYA), *Tursiops truncatus*-*Monodelphis domestica* divergence (124.6–138.4 MYA), and *Ornithorhynchus anatinus*-*Tursiops truncatus* divergence (162.5–191.1 MYA) were derived from a previously published paper<sup>7</sup>.

The mutation rate (substitutions per site per year) was calculated based on the maximum likelihood (ML) tree distance and split time estimates. The maximum likelihood phylogenetic tree and split times were computed for the four-fold degenerate sites of the concatenated coding sequences (CDS) of single-copy gene families.

## Overexpression of glutathione and glutathione disulfide

The unique amino acid changes and positive selection in the glutathione metabolism associated genes may have been an evolutionary outcome of increased resistance to hypoxic conditions during diving. The over-representation of glutathione (GSH) and glutathione disulfide (GSSG) were validated in kidney Sp1k cells of Atlantic spotted dolphin (*Stenella frontalis*), which is the only available whale cell lines. GSH and GSSG were more highly induced in Atlantic spotted dolphin cells than in human kidney HEK293 cells under hypoxic conditions (low O<sub>2</sub>); i.e., the glutathione/glutathione disulfide ratio was extremely reduced in the Atlantic spotted dolphin kidney cells. GSH and GSSG were also highly unregulated in Atlantic spotted dolphin kidney cells upon hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>) treatment. These results provide strong evidence for the adaptation of whales to hypoxic conditions during diving.

### Cell culture and treatment of cells with H<sub>2</sub>O<sub>2</sub>

HEK293 (human kidney) cells were maintained in Dulbecco's modified Eagle's medium (DMEM; Gibco) supplemented with 10% fetal bovine serum, whereas Sp1k (Atlantic spotted dolphin kidney) cells were maintained in Eagle's minimum essential medium (EMEM) with Earle's balanced salt solution (EBSS) supplemented with 2 mM glutamine, 2% non-Essential Amino Acids (NEAA), and 10% fetal bovine serum at 37°C in a humidified incubator with 5% CO<sub>2</sub>. HEK293 cells and Sp1k cells were obtained from the American Type Culture Collection (Rockville, MD) and the Health Protection Agency (Salisbury, SP), respectively. The cells were treated with 100 μM H<sub>2</sub>O<sub>2</sub> for 1 h.

### Hypoxic incubation and estimation of glutathione levels

HEK293 and Sp1k cells were placed in a hypoxic incubator (Hela Cell 150, Thermo Scientific Inc.) for 6 h or 12 h or were maintained in the tissue culture incubator for 5 min. The cells were lysed using total or oxidized glutathione reagent, and the glutathione (GSH) and glutathione disulfide (GSSG) contents of the HEK293 and Sp1k cells were measured using a GSH/GSSG-Glo<sup>TM</sup> Assay Kit (Promega, Madison, WI) according to the manufacturer's instructions.

## Rapid evolution

We evaluated the  $d_N/d_S$  ( $Ka/Ks, w$ ) ratio to identify Gene Ontology (GO) categories that were significantly over-represented or under-represented in the minke whale genome<sup>8</sup>. First, we selected single copy orthologs from the minke whale, bottlenose dolphin, and human, which is an outgroup. Second, PRANK<sup>9</sup> was used to generate multiple sequence alignments of the single copy orthologs, and the alignments were filtered using Gblocks<sup>10</sup>. The  $d_N$  (nonsynonymous substitutions per nonsynonymous site) and  $d_S$  (synonymous substitutions per synonymous site) were estimated by PAML<sup>6</sup> using an F3x4 codon frequency model and the REV substitution matrix. Human GO annotations from the Ensemble database (release 69) were used to analyze the GO categories of rapidly or slowly evolving genes. The  $d_N$  and  $d_S$  values for each GO category were calculated, and any GO category containing fewer than 20 orthologs was filtered out. For a given GO category, the probability of observing an equal or greater number of non-synonymous substitutions was calculated assuming a binomial distribution.

To determine whether a subset of the GO categories was evolving under significantly high (or low) constraints, we repeated this procedure 10,000 times on the same dataset after randomly permuting the GO annotations to test whether the probability was less than a threshold value. Then, we determined the GO categories with the P-values of less than 0.01. Finally, 25 and 9 GO categories were selected as rapidly and slowly evolving, respectively.

## Copy number variations in the fin whale and finless porpoise

The copy number differences (CND) in the fin whale and finless porpoise (target group) were calculated by comparing their gene copy numbers to those in the reference genomes (reference group: minke whale and bottlenose dolphin, respectively). To calculate the CND, short reads of the target group species were mapped to the reference genomes. Short reads of the reference group species were also mapped to themselves using BWA-0.6.2<sup>11</sup> to determine the gene copy number of the reference group species. The copy number was calculated using readDepth<sup>12</sup>. Genes with a conserved copy number were selected and used for the CND calculation under the following criteria: 1) level threshold: the gene copy number of both genomes (target and reference groups) is  $\times 2N$ ; 2) the level threshold: gene copy number of the reference group species is  $\geq 2N$ ; 3) consistency: the frequency of the same gene copy number is over 90 %.

### **Repeat annotation**

Repetitive elements constitute 37.3% of the minke whale reference genome. The genome was searched for tandem repeats using Tandem Repeats Finder<sup>13</sup> (version 4.04). Transposable elements (TEs) were identified using homology-based approaches. A homology-based approach was employed using Repbase<sup>14</sup> (version 16.10), a widely used database of known repeats, and a *de novo* repeat library generated by RepeatModeler<sup>15</sup>. This database was used to find repeats with software programs such as RepeatMasker<sup>16</sup> (version 3.3.0).

TEs were also detected using mammalian repeat elements alone, and 31.3% of all TEs examined were detected in the minke whale genome. The compositions of the TEs in the minke whale genome were determined by *de novo* RNA sequencing.

### **Non-coding RNAs**

Four non-coding RNAs (ncRNAs) types were annotated: microRNAs (miRNAs), transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), and small nuclear RNAs (snRNAs). These four non-coding RNA types were detected in the minke whale genome by searching databases using the complete genome sequence. tRNAscan-SE<sup>17</sup> (version 1.23) was performed using a SINE premasked genome to search for reliable tRNA positions. snRNAs and miRNAs were sought using a two-step method: after aligning with BLAST, INFERNAL was used to search for putative sequences in the Rfam database<sup>18</sup> (release 9.1). The genome was searched with human full-length rRNAs as queries for possible rRNA predictions in the minke whale genome.

## Whole genome structure

Whole genome alignments were used for comparative analyses to identify and characterize functional elements. Differences in genome structure were detected using LASTZ<sup>19</sup> by aligning the minke whale scaffolds to the bottlenose dolphin (*T. truncatus*) and wild pig (*S. scrofa*) sequences after repeat masking. A total of 30~45% of the genome region was syntenic among *B. acutorostrata*, *S. scrofa*, and *T. truncatus*. Considering that approximately 40% of the genomes were repeat sequences and masked with Ns before alignment, only 15–30% of the genome sequences were species-specific, non-repeat-containing sequences. This indicates that the three genomes had a relatively high level of genomic synteny.

## Segmental duplication

Segmental duplications are duplicated blocks of genomic DNA that typically range in size from 1 to 200 Kb. These blocks often contain sequence features such as high-copy-number repeats and gene sequences with intron-exon structures. Compared with whole-genome duplication, segmental duplications may have played a more significant role in gene and genome evolution.

A self-to-self sequence alignment identified 11,701 recent segmental duplicated fragments (>90% identity, >1 Kb length) with a total length of 33.4 Mb (1.2%) in the minke whale reference genome assembly.

## Analysis of whale *Hox* genes

*Hox* gene clusters, which play an important role in many aspects of the body plan and embryonic development, are evolutionarily highly conserved in mammals<sup>20</sup>. Unfortunately, *Hox* gene clusters are not well assembled in many draft genomes, because of the high degree of sequence similarity among the gene members of the clusters. Therefore, we re-constructed each *Hox* gene using following methods.

- 1) First, we searched the assembled whale scaffolds and already available mammalian chromosomes (minke whale, bottlenose dolphin, cow, pig, human, and mouse) using tBLASTn (default options) for finding the mis- or un-predicted *Hox* genes. The *Hox* gene query sequence for the search was the one closest to the average length of the six mammalian sequences
- 2) We tried to re-assemble the minke whale scaffolds with the longer *K*-mer size (*K*=63-mer) to improve the assembly quality for the *Hox* gene cluster regions. Then, tBLASTn with default options was used to search for each *Hox* gene in the re-assembled scaffolds.
- 3) We aligned the all the minke whale reads to the cow genome (UMD3.1.68), which has near complete sets of *Hox* genes. The consensus sequence for each *Hox* gene was generated using the SAMTools-0.1.18<sup>21</sup>. The command to generate the consensus sequence was “samtools mpileup -uf ref.fa aln.bam | bcftools view -cg - | vcfutils.pl vcf2fq > cons.fq”.

To provide insights into the whales' morphological adaptations, we calculated  $d_N/d_S$  for the *Hox* genes in the minke whale and bottlenose dolphin, compared to terrestrial mammals (cow, pig, human, and mouse) using the CODEML program in PAML. The one-ratio model (M0) was used to estimate the general selective pressure acting among the six species. The free-ratios model (M1) was then used to calculate the  $d_N/d_S$  ratio for each branch. We found several *Hox* genes that were positively selected in the whale lineage. Among them, *Hoxb1* and *Hoxb2* are involved in hindbrain formation and craniofacial morphogenesis<sup>22,23</sup>. *Hoxa5* and *Hoxb5* contribute to morphogenesis and development of lung<sup>24,25</sup>. *Hoxd12* and *Hoxd13* are related to limb development<sup>26,27</sup>. Therefore, the positive selection found in some *Hox* genes is likely to be responsible for the development of the unique morphological adaptations, such as flipper, hind limb loss, and streamlined skull, which allow whales to live permanently in the aquatic environment.

## Echolocation

In bats and toothed whales, echolocation is an important, if not the primary, means of finding and capturing prey, navigating, maintaining group cohesion, and avoiding predators<sup>28</sup>. Prestin, which is a motor protein that is found in the outer hair cells of the inner ear of the mammalian cochlea, is linked directly to the evolution of high-frequency hearing in cetaceans, and the prestin sequences in these animals contain amino acid replacements related to vocalization frequency and an associated increased in auditory sensitivity at higher frequencies<sup>29</sup>. Species-specific amino acid changes were investigated in the minke whale, fin whale, bottlenose dolphin, finless porpoise, human, cow, and pig. As expected, the number of amino acid replacements was highest in the bottlenose dolphin and finless porpoise, whereas the number of amino acid replacements in the baleen whale, which does not use echolocation, was similar to that in land animals.

## Blood clotting

Blood clotting and fibrinolytic systems were studied in the plasma of a sei whale (*Balaenoptera borealis*), which showed that the whale plasma was deficient in the Hageman factor (factor XII), Fletcher factor (a plasma prekallikrein), and PTA (factor XI)<sup>30</sup>. However, it is well known that whales are resilient and have an efficient wound healing ability. An examination of genes related to blood clotting revealed that the minke whale and bottlenose dolphin genomes did not contain the Fletcher factor. The Fletcher factor deficiency in the minke whale was confirmed by DNA reads and RNA transcriptome mapping. The factor II (prothrombin), V, VIII, X, *PLAT* (tissue-type plasminogen activator), and *FGB* (fibrinogen) genes each contained whale-specific amino acid changes. In particular, substitution in factor II in the minke whale and bottlenose dolphin were predicted to cause protein functional changes. Furthermore, substitutions in the factor VIII and *FGB* genes in the minke whale and bottlenose dolphin, respectively, were predicted to cause protein functional changes. Interestingly, the zebra fish, puffer fish, and lamprey also lack several blood clotting factors, such as Fletcher factor, factor XI, and factor XII<sup>31</sup>, suggesting that a deficiency in certain blood clotting factors may be a common feature of aquatic animals.

### **Highly expressed organ-specific genes**

Highly expressed organ-specific genes were detected using transcriptome data from the brain, heart, kidney, liver, lung, muscle01, muscle02, and muscle03 tissue samples. The highly expressed genes were selected only if their expression levels were greater than two times the fragments per kilobase of transcript per million mapped reads (FPKM) compared with that of other organs. Genes with an FPKM <10.0 were discarded.

RNA-seq analysis was conducted with TopHat<sup>32</sup> and Cufflinks<sup>33</sup> using the following methods<sup>34</sup>:

- 1) Read filtering
  - Reads in which ambiguous bases (“N”) comprised >10% of bases
  - Reads with an average quality of <Q20
  - Reads in which nucleotides of <Q20 comprised >40% of the reads
- 2) Reads were mapped to minke whale scaffolds using TopHat.
- 3) Reads were assembled into transcripts using Cufflinks.
- 4) Genes and transcripts for the aligned reads were reported using Cuffdiff, which is part of the Cufflinks package.

### **Neuroglobin and cytoglobin expression levels**

Neuroglobin increases the oxygen influx to the central and peripheral nervous system (the brain) where it provides protection under hypoxic conditions. It is expressed mainly in the brain and other nervous tissues in vertebrates. Neuroglobins are more likely to bring in nitric oxide to protect neuronal survival and recovery in areas where the oxygen supply is reduced<sup>35</sup>. The expression levels of neuroglobin and cytoglobin were analyzed using RNA-seq data, which showed that neuroglobin and cytoglobin were more highly expressed in the brain than in other organs, as reported previously<sup>36</sup>.

## Supplementary References

1. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* **25**, 402-408 (2001).
2. Niimura, Y. & Nei, M. Evolutionary dynamics of olfactory and other chemosensory receptor genes in vertebrates. *J. Hum. Genet.* **51**, 505-517, (2006).
3. Nguyen, D. T. et al. The complete swine olfactory subgenome: expansion of the olfactory gene repertoire in the pig genome. *BMC Genomics* **13**, 584 (2012).
4. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59**, 307-321 (2010).
5. MCMCTREE at <http://abacus.gene.ucl.ac.uk/software/paml.html>
6. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586-1591 (2007).
7. Benton, M. J. & Donoghue, P. C. Paleontological evidence to date the tree of life. *Mol. Biol. Evol.* **24**, 26-53 (2007).
8. Qiu, Q. *et al.* The yak genome and adaptation to life at high altitude. *Nat. Genet.* **44**, 946-949 (2012).
9. Löytynoja, A. & Goldman, N. An algorithm for progressive multiple alignment of sequences with insertions. *Proc. Natl. Acad. Sci. USA* **102**, 10557-10562 (2005).
10. Talavera, G. & Castresana, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **56**, 564-577 (2007).
11. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
12. Miller, C. A., Hampton, O., Coarfa, C. & Milosavljevic, A. ReadDepth: a parallel R package for detecting copy number alterations from short sequencing reads. *PLoS One* **6**, e16327 (2011).
13. Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* **27**, 573-580 (1999).
14. Repbase Release 16.02 at <http://www.girinst.org/replbase/index.html>
15. Price, A. L., Jones, N. C. & Pevzner, P. A. De novo identification of repeat families in large genomes. *Bioinformatics* **21**, i351-i358 (2005).
16. Smit, A. F. A., Hubley, R. & Green, P. RepeatMasker at <http://repeatmasker.org>
17. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer

RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955-964 (1997).

18. Griffiths-Jones, S. *et al.* Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res.* **33**, D121-124 (2005).

19. LASTZ at [http://www.bx.psu.edu/miller\\_lab/dist/README.lastz-1.02.00/README.lastz-1.02.00a.html](http://www.bx.psu.edu/miller_lab/dist/README.lastz-1.02.00/README.lastz-1.02.00a.html)

20. Liang, L. *et al.* Adaptive evolution of the Hox gene family for development in bats and dolphins. *PLoS One* **8**, e65944 (2013).

21. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).

22. Tümpel, S., Wiedemann, L. M. & Krumlauf, R. Hox genes and segmentation of the vertebrate hindbrain. *Curr. Top. Dev. Biol.* **88**, 103-137 (2009).

23. Yan, Y. L., Jowett, T. & Postlethwait, J. H. Ectopic expression of *hoxb2* after retinoic acid treatment or mRNA injection: disruption of hindbrain and craniofacial morphogenesis in zebrafish embryos. *Dev. Dyn.* **213**, 370-385 (1998).

24. Boucherat, O. *et al.* Partial functional redundancy between *Hoxa5* and *Hoxb5* paralog genes during lung morphogenesis. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **304**, L817-830 (2013).

25. Mujahid, S., Nielsen, H. C. & Volpe, M. V. MiR-221 and miR-130a regulate lung airway and vascular development. *PLoS One* **8**, e55911 (2013).

26. Davis, A. P. & Capecchi, M. R. A mutational analysis of the 5' *HoxD* genes: dissection of genetic interactions during limb development in the mouse. *Development* **122**, 1175-1185 (1996).

27. Zákány, J., Kmita, M. & Duboule, D. A dual role for Hox genes in limb anterior-posterior asymmetry. *Science* **304**, 1699-1672 (2004).

28. Branstetter, B. K., Finneran, J.J., Fletcher, E.A., Weisman, B.C. & Ridgway, S.H. Dolphins can maintain vigilant behavior through echolocation for 15 days without interruption or cognitive impairment. *PLoS One* **7**, e47478 (2012).

29. Liu, Y., Rossiter, S. J., Han, X., Cotton, J.A. & Zhang, S. Cetaceans on a molecular fast track to ultrasonic hearing. *Curr. Biol.* **20**, 1834-1839 (2010).

30. Saito, H., Poon, M., Goldsmith, G. H., Ratnoff, O. D. & Arnason, U. Studies on the blood clotting and fibrinolytic system in the plasma from a sei (baleen) whale. *Proc. Soc. Exp. Biol. Med.* **152**, 503-507 (1976).

31. Ponczek, M. B., Gailani, D. & Doolittle, R. F. Evolution of the contact phase of vertebrate blood coagulation. *J. Thromb. Haemost.* **6**, 1876-1883 (2008).

32. Trapnell, C., Pachter, L. & Salzberg, S.L. TopHat: discovering splice junctions with

RNA-seq. *Bioinformatics* **25**, 1105-1111 (2009).

33. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511-515 (2010).

34. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562-578 (2012).

35. Burmester, T., Weich, B., Reinhardt, S. & Hankeln, T. A vertebrate globin expressed in the brain. *Nature* **407**, 520-523 (2000).

36. Schneuer, M. *et al.* Neuroglobin of seals and whales: evidence for a divergent role in the diving brain. *Neuroscience* **223**, 35-44 (2012).